Vol. XIII DECEMBER, 1929 No. 6

The Journal of APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University, Athens, Ohio
WITH A BOARD OF CO-OPERATING EDITORS

TABLE OF CONTENTS

Charles Scott Berry and Clara B. Stoddard. An Experiment with Lispers 543
Donald R. Gorham and R. A. Brotemarkle. Challenging Three Standardized Emotional Tests for Validity and Employability 554
T. A. Langlie. What is Measured by the Iowa "Aptitude" Tests? 589
Eleanor Olmstead Miller. A Study of the Preschool Child's Picture and Story Books by the Battery of Tests Method 592
Howard T. Hovde. The Relative Effects of Size of Type, Leading and Context 600
Ethel Moriwaki. Validities of Judgments of Intelligence 630
Starke Rosecrans Hathaway. A Comparative Study of Psychogalvanic and Association Time Measures—A New Psychogalvanic Apparatus 632
News and Notes 647
Book Reviews 648
New Books and Pamphlets 652

PUBLISHED BI-MONTHLY BY THE EDITOR
All communications should be addressed to the Editor,
Ohio University, Athens, Ohio
Printed by
WAVERLY PRESS, INC.
Baltimore, Md.

Entered as second class matter June 3, 1924, at the post office at Baltimore, Maryland, under the act of March 3, 1879.
Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922.

Copyright 1929, by James P. Porter

The Journal of Applied Psychology

CO-OPERATING EDITORS

W. F. Book, Indiana University; W. V. Bingham, Personnel Research Federation, New York; June E. Downey, University of Wyoming; Robert H. Gault, Northwestern University; William Healy, Baker Foundation, Juvenile Court, Boston; H. L. Hollingworth, Columbia University; Joseph Jastrow, University of Wisconsin; Herbert S. Langfeld, Princeton University; James H. Leuba, Bryn Mawr College; Henry C. McComas, Princeton University; Walter D. Scott, Northwestern University; C. E. Seashore, University of Iowa; Daniel Starch, Boston, Mass.; Edward K. Strong, Jr., Leland Stanford Jr. University; L. M. Terman, Leland Stanford Jr. University; E. L. Thorndike, Columbia University; J. E. Wallace Wallin, Department of Education, Baltimore, Md.; G. M. Whipple, Danvers, Mass.; Helen Thompson Woolley, Teachers College, Columbia University.

The scope of The Journal of Applied Psychology will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this Journal; in every case the year, volume and pages should be specified.

The Journal of Applied Psychology is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

The Journal of Applied Psychology

Edited by
JAMES P. PORTER
Ohio University
With a Board of Co-operating Editors

VOLUME XIII
1929

PUBLISHED BY THE EDITORS
WAVERLY PRESS, INC.
MOUNT ROYAL AND GUILFORD AVENUES
BALTIMORE, MD.

CONTENTS

E. J. Asher: The Predictive Value of Mental Tests that Satisfy Spearman's Tetrad Criterion 152

Robert Jaudon Ball: An Objective Measure of Emotional Instability 226

J. E. Bathurst and Norma V. Scheidemann: The College Psychology Test 81

C. E. Benson: See Lucas and Benson.

C. E. Benson and D. B. Lucas: The Historical Trend of Negative Appeals in Advertising 346

Charles Scott Berry and Clara B. Stoddard: An Experiment with Lispers 543

William F. Book: How a Special Disability in Spelling was Diagnosed and Corrected 378

M. Thelma Bouton: *See* Broom, Thompson and Bouton.

M. Eustace Broom, Blanche Thompson and M. Thelma Bouton: Sex Difference in Handwriting 159

R. A. Brotemarkle: *See* Gillespie and Brotemarkle.

R. A. Brotemarkle: *See* Gorham and Brotemarkle.

W. F. Cline: *See* Stoke and Cline.

Josephine Douglas and J. W. Lawson: Measurement of Reading Skills in Ability Groups 494

Harl R. Douglass and C. L. Huffaker: Correlation between Intelligence Quotient and Accomplishment Quotient 76

Karl C. Garrison: An Investigation of Some Simple Speed Activities 167

F. Rodney Gillespie and R. A. Brotemarkle: Interpolated Revision of College Adult-Level Norms for Personnel Procedure 325

Donald R. Gorham and R. A. Brotemarkle: Challenging Three Standardized Tests for Validity and Employability 554

Starke R. Hathaway: A Comparative Study of Psychogalvanic and Association Time Measures—A New Psychogalvanic 632

M. H. Heim: *See* Newhall and Heim.

J. A. Highsmith: Selecting Musical Talent 486

Howard T. Hovde: The Relative Effects of Size of Type, Leading and Context 600

Hsiao Hung Hsiao: The Mentality of the Chinese and Japanese 9

C. L. HUFFAKER: *See* Douglass and Huffaker.
T. A. JACKSON: Errors in Self Judgment 372
FRIEDA A. KIEFER: Manual Motor Correlation in Superior Children 357
F. KUHLMANN: The Pearson r and a Further Note on the Kuhlmann-Anderson Tests 32
T. A. LANGLIE: What is Measured by the Iowa "Aptitude" Tests? ... 589
HOWARD P. LONGSTAFF AND JAMES P. PORTER: Objective Measures of Ability and Achievement in General Psychology 173
D. B. LUCAS AND C. E. BENSON: The Relative Values of Positive and Negative Advertising Appeals as Measured by Coupons Returned 274
D. B. LUCAS: *See* Benson and Lucas.
J. H. MCFADDEN: A Further Note on the Differential IQ's of Siblings 86
ELEANOR OLMSTEAD MILLER: A Study of the Preschool Child's Picture and Story Books by the Battery of Tests Method 592
ALAN H. MONROE: Effect of Bodily Action on Voice Intensity 516
RUTH P. MONTGOMERY: *See* Poull and Montgomery.
ETHEL MORIWAKI: Validities of Judgments of Intelligence 630
YALE S. NATHANSON: A Conceptual Basis of Habit Modification 469
S. M. NEWHALL AND M. H. HEIM: Memory Value of Absolute Size in Magazine Advertising 62
MARGARET OTIS: Improvement of Feeble-Minded Girls Over Sixteen Years of Age 301
DONALD G. PATTERSON: *See* Tinker and Paterson.
RUDOLF PINTNER: Speech and Speech-Reading Tests for the Deaf ... 220
JAMES P. PORTER: *See* Longstaff and Porter.
S. D. PORTEUS: The Social Psychology of the Australian Aboriginal 131
LOUISE E. POULL AND RUTH P. MONTGOMERY: The Porteus Maze Test as a Discriminative Measure in Delinquency 145
H. H. REMMERS: The Measurement of Interest differences between Students of Engineering and of Agriculture 105
FRANK S. SALISBURY AND H. B. SMITH: Prognosis of Sight Singing Ability 425
PAUL V. SANGREN: Comparative Validity of Primary Intelligence Tests 394
NORMA V. SCHEIDEMANN: An Experiment in Teaching Psychology .. 188
NORMA V. SCHEIDEMANN: *See* Bathurst and Scheidemann.
RICHARD S. SCHULTZ: A Mental Alertness Examination for the Working Age Level 46
KENYON J. SCUDDER: The Predictive Value of General Intelligence Tests in the Selection of Junior Accountants and Bookkeepers ... 1

H. B. Smith: *See* Salisbury and Smith.
Melissa Branson Stedman: A Study of the Possibility of Prognosis of School Success in Typewriting..... 505
Clara B. Stoddard: *See* Berry and Stoddard.
Emily Leatherman Stogdill: The Maladjusted College Student—A Further Study with Results..... 440
Stuart M. Stoke and W. F. Cline: The Avocations of One Hundred College Freshmen..... 257
Blanche Thompson: *See* Broom, Thompson and Bouton.
Miles A. Tinker and Donald G. Paterson: Studies of Typographical Factors Influencing Speed of Reading. II. Size of Type..... 120
Miles A. Tinker and Donald G. Paterson: Studies of Typographical Factors Influencing Speed of Reading. III. Length of Line..... 205
Chr. Hansen-Tybjerg: Differences in the Talents and Abilities of Young People..... 451
Roswell H. Whitman: A Short Scale for Measuring Introversion-Extroversion..... 499
Margaret M. Wood: Mental Test Findings with Armenian, Turkish, Greek, and Bulgarian Subjects..... 266

AN EXPERIMENT WITH LISPERS

CHARLES SCOTT BERRY

Professor of Educational Psychology, University of Michigan, and Consultant in Special Education, Detroit Public Schools

AND

CLARA B. STODDARD

Supervisor of Speech Improvement, Detroit Public Schools

This experiment was undertaken to determine to what extent lispers who receive no corrective speech training improve as compared with those who receive such training.

During the last of September and first of October 448 lispers in the regular grades of the Detroit public schools were tested by 28 speech improvement teachers. The lispers in grades one, two and three were given the Detroit Articulation Test I, Reading Form; and those in grades seven to twelve, the Detroit Articulation Tests II and III.[1]

After the lispers had been tested they were paired as far as possible according to school, grade, extent of defect, intelligence, age and home language. From the middle of October to the first of June one member of each of the 224 pairs of lispers was given two lessons per week of 75 minutes duration in the regular speech improvement classes. The other member of each pair received no corrective speech training. The lispers who received the lessons in speech improvement are known as the "experimental group," and those who received no lessons in speech training as the "control group." In June the members of both the experimental and the control groups were retested by the same speech improvement teachers who tested them in the fall. In fact, this experiment consisted of twenty-eight independent experiments, as each one of the twenty-eight

[1] These tests are published by the Detroit Board of Education.

teachers selected and tested sixteen lispers, paired them, taught her own experimental group in her own speech improvement classes and in June retested her own experimental and control groups.

TABLE 1

Comparison of the experimental and control groups in respect to extent of defect, grade distribution, intelligence, age, home language and sex

	EXPERIMENTAL GROUP	CONTROL GROUP
Number of cases	172	172
Extent of defect:		
Average	55.2	55.7
Medium	56.0	56.5
25 Percentile	49.5	51.0
75 Percentile	62.0	62.0
Quartile Deviation	6.2	5.5
	per cent	*per cent*
Grade Distribution:		
Grades 1, 2, 3	70.5	70.5
Grades 4, 5, 6	27.1	27.1
Grades 7, 8	2.4	2.4
Intelligence		
X	11.0	10.5
Y	55.8	59.3
Z	33.2	30.2
Age:		
5, 6, 7	43.6	45.9
8, 9, 10, +	56.4	54.1
Home Language:		
English	56.4	58.1
Foreign	43.6	41.9
Sex:		
Boys	65.7	60.5
Girls	34.3	39.5

Of the 224 pairs of lispers with whom the 28 teachers started in October there remained in June 172 pairs for whom complete data were available. Only the results for these 172 pairs are presented in the following tables.

In table 1 the experimental and control groups are compared

in respect to extent of defect, grade distribution, intelligence, age, home language, and sex.

In the articulation tests used to determine the extent of defect, the possible scores range from 100 to zero, 100 representing speech free from errors of voice and articulation. In these tests the score is determined by subtracting the number of errors from 100. In table 1 the extent of defect is represented by the scores made in the articulation test given at the beginning of the experiment. Not only are the experimental and control groups closely similar in average and median scores but also in variability as expressed by the quartile deviation.

In the Detroit Public Schools pupils are classified by intelligence tests into three groups known as X, Y, and Z. In the X group are found the 20 per cent that test highest in a given age group; the next 60 per cent make up the Y group and the lowest 20 per cent comprise the Z group. From table 1 it is evident that there is no great difference between the experimental and control groups in respect to intelligence. Neither is there any significant difference between the two groups in respect to grade distribution, age, and language spoken at home. The greatest difference is in the distribution of boys and girls, there being only 34 per cent of girls in the experimental group as compared with 39 per cent in the control group. However, this proved to be to the advantage of the control group as the girls were found to make a slightly greater gain than the boys.

In table 2 the facts relative to the improvement made by the experimental and control groups from October to June are presented.

The improvement in speech in these 172 pairs of lispers was determined by subtracting algebraically the October score in the articulation test from the June score in the same test. The remainder thus obtained represented the difference in the number of errors made in the two tests. Thus gain or improvement is measured by the decrease in the number of errors. In table 2 the average improvement for the experimental group is seen to be 27.5 as compared with 4.3 for the control group. This

means that the experimental group averaged 27.5 less errors in the second test than in the first and that the control group averaged 4.3 less errors in the second test than in the first. The median improvement of the experimental group is 27 points greater than that of the control group. The marked variability of improvement in the middle 50 per cent of the experimental group is indicated by the large quartile deviation of 10.2.

Some significant facts not presented in this table are that only 1.7 per cent of the experimental group failed to gain while 45.3 per cent of the control group made no gain, that only 2.3 per cent of the control group reached or exceeded the median

TABLE 2

Improvement in experimental and control groups

	EXPERIMENTAL GROUP						CONTROL GROUP					
	Average	Medium	25 Percentile	75 Percentile	Quartile Dev.	Range	Average	Medium	25 Percentile	75 Percentile	Quartile Dev.	Range
Improvement.....	27.5	29	17	37.5	10.2	−6 to 73	4.3	2	0	7	3.5	−22 to 46

improvement (29) of the experimental group, and that only 6.4 per cent of the control group exceeded the 25 percentile point of the experimental group. Furthermore, only 1.7 per cent of the control group made a score of 85 or above in the June articulation test as compared with 54.6 per cent of the experimental group. This is especially significant since pupils are usually discharged from the speech improvement classes on trial who make a score of 85 or above. In other words, more than half of the pupils in the experimental group were tentatively discharged from the speech improvement classes as no longer in need of special corrective treatment as compared with 1.7 per cent in the control group. In this connection it might

be said that the amount of improvement made by the experimental group was about the same as that made in the regular speech improvement classes during the previous year.

In connection with an experiment of this kind the question arises as to the extent to which factors other than corrective speech training may condition improvement. For example, what may be the effect of extent of defect, intelligence, age and home language on speech improvement?

In table 3 the 172 pairs of lispers in the experimental and control groups are classified according to extent of defect.

TABLE 3

Improvement in relation to extent of defect

EXTENT OF DEFECT	EXPERIMENTAL GROUP							CONTROL GROUP						
	Number of cases	Per cent of cases	Av. Imp.	Med. Imp.	25 Percentile	75 Percentile	Quartile Dev.	Number of cases	Per cent of cases	Av. Imp.	Med. Imp.	25 Percentile	75 Percentile	Quartile Dev.
Class I 49 or less	43	25.0	35.3	36	23	44	10.5	35	20.4	10.0	7	0	14	7.0
Class II 50 to 59	67	39.0	29.3	33	20	38	9.0	66	38.4	3.6	1	0	5	2.5
Class III 60 or more	62	36.0	20.2	20	10	30	10.0	71	41.2	2.1	0	0	5	2.5
Totals....	172	100.0	27.5	29	17	37.5	10.2	172	100.0	4.3	2	0	7	3.5

In Class I are those who made a score of 49 or less in the October test; in Class II, those who made a score from 50 to 59; and in Class III, those whose score was 60 or above. By referring to this table it is seen that there is a marked correlation between extent of defect and amount of improvement, the greater the extent of defect the greater the improvement. For Classes I, II, and III the median improvement is respectively 36, 33 and 20 in the experimental group, and 7, 1, and 0 in the control group. The same order of improvement is maintained in the 25 and 75 percentiles of all three classes in the

experimental group. However, in both the experimental and the control groups the three classes overlap to a marked degree in the amount of improvement.

The significant fact revealed by this table is the marked positive correlation between extent of defect and amount of improvement.

In table 4 the lispers are classified according to intelligence. This table shows that in both the experimental and control groups the lispers of Y intelligence made a greater median and average gain than those of Z or X intelligence. However, the greater gain made by the Y lispers might be due to there being

TABLE 4

Improvement in relation to intelligence

INTELLIGENCE	EXPERIMENTAL GROUP							CONTROL GROUP						
	Number of cases	Per cent of cases	Av. Imp.	Med. Imp.	25 Percentile	75 Percentile	Quartile Dev.	Number of cases	Per cent of cases	Av. Imp.	Med. Imp.	25 Percentile	75 Percentile	Quartile Dev.
X	19	11.1	26.2	28	11	39	14	18	10.5	2.8	0	0	0	0
Y	96	55.8	29.5	31.5	19	39	10	102	59.3	4.6	2	0	8	4
Z	57	33.1	24.8	24	16	35	9.5	52	30.2	4.3	1.5	0	6	3
Totals....	172	100.0	27.5	29	17	37.5	10.2	172	100.0	4.3	2	0	7	3.5

in that group a larger percentage of cases in Class I than in the groups of X or Z intelligence, for we have seen that there is a positive correlation between extent of defect and amount of gain.

In table 5 the lispers of X, Y and Z intelligence are divided according to extent of defect into Classes I, II, and III. From this table it can be seen that in the experimental group the Y's have a larger percentage of cases in Class I than have the Z's; but this is compensated for by the larger percentage of Y's in Class III. Furthermore, when we compare the median and average improvement of the Y pupils in Classes I, II, and III with the Z pupils in the corresponding classes we find that

in each class the Y pupils made a greater average and median improvement than the Z pupils.

The Y's have a larger percentage of cases in Class I, but a smaller percentage in Class III than have the X's. If the distribution of X and Y pupils into Classes I, II, and III were the

TABLE 5

Improvement in relation to intelligence in lispers similar in extent of defect

INTELLIGENCE		EXTENT OF DEFECT							
		Experimental group				Control group			
		Class I 49 or less	Class II 50 to 59	Class III 60 or more	Totals	Class I 49 or less	Class II 50 to 59	Class III 60 or more	Totals
X	Number	3	7	9	19	3	6	9	18
	Per cent	16	37	47	100	17	33	50	100
	Av. Imp.	47.3	24.5	20.4	26.2	20.5	−0.7	−0.9	2.8
	Med. Imp.	52	23	25	28	15	0	0	0
Y	Number	29	31	36	96	22	33	47	102
	Per cent	30	32	38	100	22	32	46	100
	Av. Imp.	35.8	33.4	21.1	29.5	10.6	2.6	3.0	4.6
	Med. Imp.	39	35	22.5	31.5	10	1	2	2
Z	Number	11	29	17	57	10	27	15	52
	Per cent	19	51	30	100	19	52	29	100
	Av. Imp.	30.9	26.0	18.8	24.8	5.4	5.4	1.5	4.3
	Med. Imp.	30	[30]	19	24	2.5	4	0	1.5
Totals	Number	43	67	62	172	35	66	71	172
	Per cent	25	39	36	100	20	39	41	100
	Av. Imp.	35.3	29.3	20.2	27.5	10.0	3.6	2.1	4.3
	Med. Imp.	36	33	20	29	7	1	0	2

same as in the entire experimental group, the average improvement of the X group would be approximately the same as that of the Y group.

In the control group we have little evidence of any significant relation between improvement and intelligence.

However, the facts presented in table 5 show that on the

average Z lispers make less progress in speech improvement classes than the lispers of higher intelligence.

In table 6 the lispers are divided into two groups according to age at last birthday. In one group are those 5, 6, or 7 years of age; in the other, those 8, 9, 10 or above. By reference to this table it can be seen that in the experimental group the median improvement of the younger pupils is 3 points higher than that of the older pupils. However, if the percentage distribution of the age groups in respect to extent of defect and

TABLE 6

Improvement in relation to age

	EXPERIMENTAL GROUP		CONTROL GROUP	
	Ages 5, 6, 7	Ages 8, 9, 10 or more	Ages 5, 6, 7	Ages 8, 9, 10 or more
Number of cases	75	97	79	93
Extent of defect:				
Per cent Class I—49 or less	29	22	23	18
Per cent Class II—50 to 59	35	42	38	39
Per cent Class III—60 or more	36	36	39	43
Intelligence:				
Per cent X's	17	6	13	9
Per cent Y's	61	52	69	50
Per cent Z's	22	42	18	41
Med. improvement	31.0	28.0	4.0	0.0
Av. improvement	29.6	25.9	6.6	2.4

intelligence were the same it would reduce this difference from 3 points to about 2. Also in the control group, if the percentage distribution of both age groups were the same in respect to extent of defect and intelligence, the difference in the median improvement of the two age groups would be slightly less.

If the percentage distribution were equalized as suggested still the median improvement of the younger children in both experimental and control groups would be greater than that of the older children, but not enough greater in view of the wide variability, to have much significance.

In table 7 the lispers are divided according to home language. In the "English" group are those who heard only English spoken at home, while in the "foreign" group are those who heard only a foreign language or a foreign language and English spoken at home.

It can be seen from this table that in both the experimental and control groups the lispers who heard only English at home made a greater median and average improvement than those who heard a foreign language at home. However, if the per-

TABLE 7

Improvement in relation to home language

	HOME LANGUAGE			
	Experimental group		Control group	
	English	Foreign	English	Foreign
Number of cases	97	75	100	72
Extent of defect:				
Per cent Class I—49 or less	26	24	17	25
Per cent Class II—50 to 59	37	41	37	40
Per cent Class III—60 or more	37	35	46	35
Intelligence:				
Per cent X's	16	4	13	7
Per cent Y's	59	52	64	53
Per cent Z's	25	44	23	40
Improvement Med.	31	27	2.0	0.0
Improvement Av.	28.8	25.9	4.4	4.2

centage distribution in respect to degree of defect and intelligence were the same for the two language groups, the difference in the average improvement would be decreased in the experimental group, but increased in the control group. But in neither case would the change amount to as much as one point.

The data presented in this table seem to show, other things being equal, that lispers who heard only English spoken at home made a slightly greater improvement than those who heard a foreign language at home.

The data for lispers tabulated separately for boys and girls showed no marked sex differences in the amount of improvement, although the girls made a slightly greater gain than the boys.

In this experiment speech improvement in lispers of the like extent of defect was found to be conditioned only to a slight extent by differences in intelligence, age, home language or sex as compared with corrective training in speech improvement classes.

SUMMARY AND CONCLUSIONS

In September and October 344 lispers in the public schools of Detroit were given speech tests and then paired in respect to school, grade, extent of defect, intelligence and home language.

One hundred seventy-two of these lispers, the experimental group, were given two lessons per week in the regular speech improvement classes. The remaining 172 lispers, the control group, received no lessons in corrective speech training.

In June both groups were again given speech tests. The change in speech was measured by the difference in the scores made in the two tests. The score was determined by subtracting the number of errors made in the test from one hundred. Thus improvement was measured by decrease in number of errors.

The median improvement for the experimental and control groups was respectively 29 and 2. Only 1.7 per cent of the experimental group failed to gain; but 45.3 per cent of the control group made no gain. Only 2.3 per cent of the control group reached or exceeded the median improvement of the experimental group.

In both the experimental and control groups positive correlation was found between extent of defect and amount of improvement.

The lispers of low intelligence made a smaller average improvement than those of higher intelligence.

In both the experimental and the control groups the lispers from homes where a foreign language was spoken made a smaller

average gain than those from homes where only English was spoken.

The girls made a slightly greater average improvement than the boys, but not enough greater to be significant.

In this experiment speech improvement in lispers of like extent of defect was found to be conditioned only to a slight degree by differences in intelligence, age, home language or sex, as compared with corrective training in speech improvement classes.

CHALLENGING THREE STANDARDIZED EMOTIONAL TESTS FOR VALIDITY AND EMPLOYABILITY

DONALD R. GORHAM

Eastern University

AND

R. A. BROTEMARKLE

University of Pennsylvania

Three tests stand out as conspicuous attempts to analyze the complex emotional makeup of the individual. These are the Downey Will-Temperament Test, the Pressey Cross-Out Tests, and the Brotemarkle Comparison Tests. Much of the criticism of these tests is due to a lack of understanding of their proper use. The fault lies partly at the door of the authors of the tests. Although Downey has set forth at some length an account of the construction and use of her test,[1] neither Pressey nor Brotemarkle has written a sufficient explanation of his tests to make them employable by any except experienced clinical psychologists. The result is that the tests have been misused and the results misinterpreted. In this study these tests are challenged as to their validity and employability, and the challenge is accepted with an attempt at vindication.

The method has been two-fold: first, to determine the validity of the tests by statistical comparison of the various scores; second, to determine their validity by using them as the basis of case studies.

The data were secured from the members of a seminar class at Eastern University—ten men and ten women, most of whom were college freshmen. In the case of the self-ratings and composite ratings on the Downey Will-Temperament Test, only ten students participated—the ones who were best fitted to rate one another.

[1] Downey: *The Will-Temperament and Its Testing.*

I. STATISTICAL COMPARISON

The Downey Will-Temperament Test

Procedure. Two types of ratings of the Will-Temperament Test were employed. The first type was secured in the usual way—by administering the Downey Group Will-Temperament Test. The second type was made up of the judgments of the individuals by themselves and others of the group. The group

TABLE 1

Correlations between Test Ratings, Self Ratings and Composite Ratings of Downey Will-Temperament Traits

WILL-TEMPERAMENT TRAITS	T S*		T C		C S	
	r	P.E.	r	P.E.	r	P.E.
A. Speed of Movement	.67	.12	.73	.10	.52	.15
B. Freedom from Load	−.18	.19	.08	.21	.60	.14
C. Flexibility	−.18	.20	.75	.09	.10	.21
D. Speed of Decision	.77	.09	.64	.12	.31	.19
E. Motor Impulsion	−.04	.21	−.27	.14	.71	.10
F. Reaction to Contradiction	.42	.17	.10	.21	.59	.14
G. Resistance to Opposition	.03	.21	−.40	.17	.61	.13
H. Finality of Judgment	.08	.21	−.05	.17	.75	.09
I. Motor Inhibition	.51	.15	.42	.17	.81	.07
J. Interest in Detail	.49	.17	.47	.19	.60	.14
K. Co-ordination of Impulses	.30	.19	.56	.14	.48	.16
L. Volitional Perseveration	.12	.21	−.06	.21	.29	.19
Mean	.249	.18	.247	.19	.531	.13

* T = Test Ratings; S = Self Ratings; C = Composite Ratings.

supplying these ratings consisted of ten students, mostly college freshmen. The author's definition of the Will-Temperament traits was used as the basis for the judgments. The composite rating of each student was secured by averaging the ratings of the other nine members of the group.

Results. The coefficients of correlation between the test, composite and self-ratings and the twelve Will-Temperament traits are shown in table 1. The data show that the average

correlation of both self and composite ratings with the test ratings was almost identical, the coefficient being about .25 ±.18. In regard to certain traits, there seemed a marked ability to judge on the part of the students, viz.:

Speed of Decision	.715
Speed of Movement	.70
Interest in Detail	.48
Motor Inhibition	.465
Co-ordination of Impulses	.43

In regard to flexibility, there was not a consistent judgment, the composite ratings giving a very high coefficient, while the self ratings gave a negative coefficient—when correlated with the test scores.

In certain traits the students showed a marked agreement among themselves as to the amount possessed by themselves and others. This is shown by the CS coefficients in table 1. In the following traits, marked agreement in rating was manifest:

Motor Inhibition	.81
Finality of Judgment	.75
Motor Impulsion	.71
Resistance to Opposition	.61
Interest in Detail	.60
Freedom from Load	.60
Reaction to Contradiction	.59
Speed of Movement	.52
Co-ordination of Impulses	.48

It appears, then, that in at least three-fourths of the traits the students evidently understood the definitions of the traits and were able to rate themselves and others on the degree to which they were possessed. Unfortunately, although they agreed among themselves, their ratings on the majority of these tests did not agree with the rating as measured by the test. Whatever the Downey test measures, we must conclude that it does not measure what this group of students observe as they understand the definitions of its component factors.

The Pressey Cross-Out Tests

Procedure. The Pressey Cross-Out Tests were given to the group. Differential scores for men and women were computed as follows: The forty words of the entire test were found which were crossed out most often by the women; also the analagous forty words crossed out most often by the men. The rating of each case equals the number crossed out of the women's words, less the number crossed out of the men's words. The results are shown graphically in chart 1. The striking sex difference as shown by words crossed out on the test is immediately apparent. When it is remembered that the words crossed out—if analyzed—furnish an index to the emotional life of the individual, these results become truly significant.[2]

Test 1 of the Pressey Cross-Out Test was analyzed for each case by determining the number of each type of words crossed out, viz., disgust, fear, sex, suspicion, and the jokers. Table 2 gives the results for each case, together with the means for men and women, and the tentative norms given by Pressey. Chart 2 shows the averages for men and women compared with the norms. Although the chart serves to compare the averages with the norms, it must be admitted that the comparisons are not significant, since one abnormal score in the group would be completely obliterated by the whole.

The same procedure was followed for Test 4, and the data are presented similarly in table 3 and chart 3.

The Brotemarkle Comparison Test

Procedure and results. The Brotemarkle Comparison Test was given to the group and the data are shown in table 4, expressed by the total deviation.

Five Test Scores Compared

After the foregoing examination of each test individually, the five sets of test scores were thrown together for the purpose of comparison. The data are presented in table 5. After all,

[2] *Journal of Applied Psychology* for April, 1921.

now that we have the scores for the twenty cases, what do they mean? Is Downey justified in stating in her Test Manual that

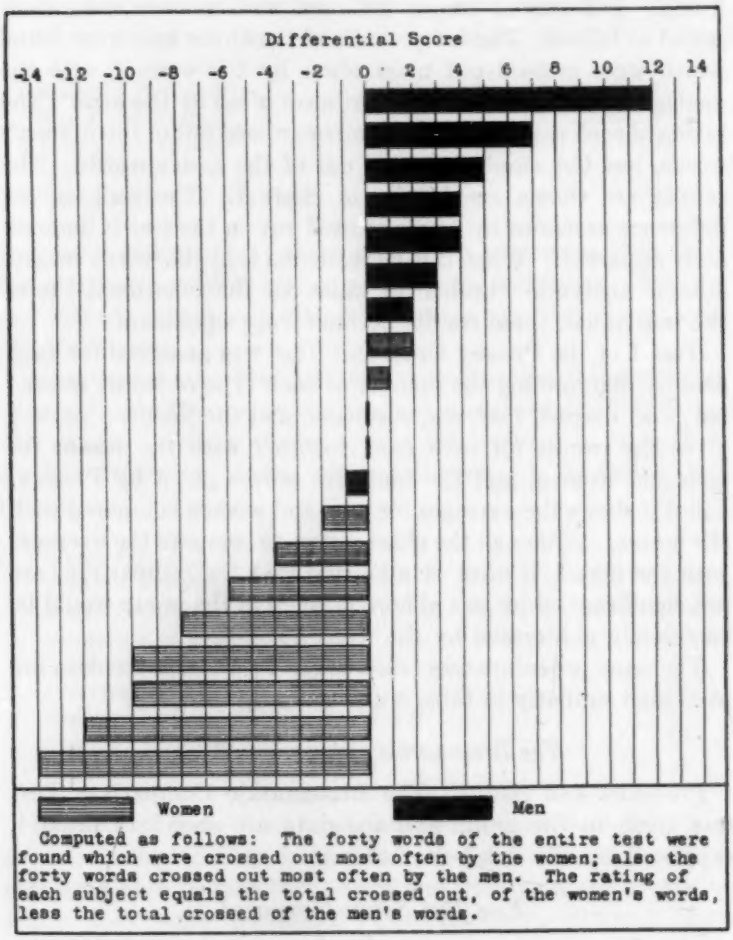

Differential Score
-14 -12 -10 -8 -6 -4 -2 0 2 4 6 8 10 12 14
Women
Men
Computed as follows: The forty words of the entire test were found which were crossed out most often by the women; also the forty words crossed out most often by the men. The rating of each subject equals the total crossed out, of the women's words, less the total crossed of the men's words.

CHART 1. DIFFERENTIAL SCORE ON PRESSEY X-O TEST

a high total score on her test indicates a strong personality while a low total score represents a weak personality? If so,

TABLE 2

Analysis of Test I—Pressey X-O Test

CASE NUMBER	DISGUST	FEAR	SEX	SUSPICION	JOKERS	TOTAL
			Women			
4	19	12	10	14		55
7	20	8	18	8		54
8	14	2	10	4		30
10	7	2	10	6		25
12	12	15	8	9		44
13	15	10	10	12		47
15	14	10	9	16		49
16	10	2	8	13		33
17	15	13	7	15		50
20	18	11	10	15		54
Mean.....	14.4	8.5	10.0	11.2		44.1
			Men			
1	16	3	3	5		27
2	21	9	11	4		45
3	14	12	9	15		50
5	11	7	8	8		34
6	16	10	12	13		51
9	8	3	1	5		17
11	8	5	3	2		18
14	18	16	11	18		63
18	13	7	6	10		36
19	16	5	8	2		31
Mean.....	14.1	7.7	7.2	8.2		37.2
Total Mean...	14.25	8.1	8.6	9.7		40.6
Norm.....	14.2	8.6	13.0	10.1	.3	46.2

Tentative norms given by Pressey in Journal of Abnormal Psychology for April, 1921.

Each column gives the words crossed out by each case with the averages for men and women, the total cases, and the Pressey norms.

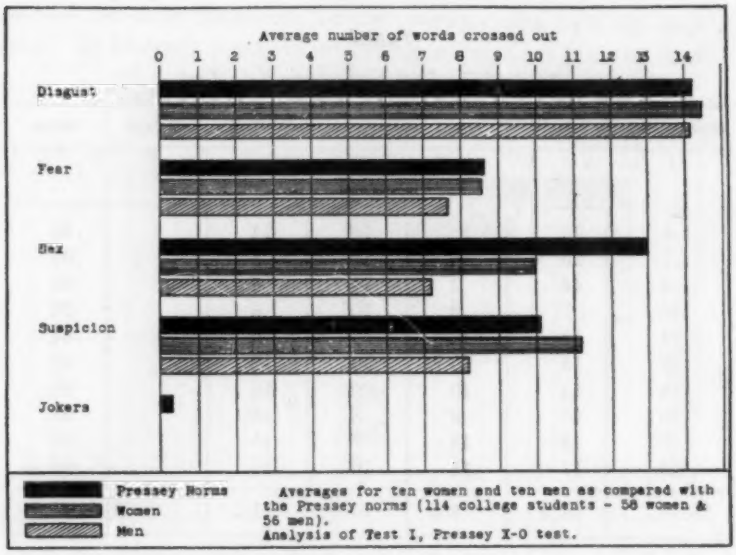

Average number of words crossed out
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14
Disgust
Fear
Sex
Suspicion
Jokers
Pressey Norms
Women
Men
Averages for ten women and ten men as compared with the Pressey norms (114 college students - 58 women & 56 men).
Analysis of Test I, Pressey X-O test.

CHART 2. ANALYSIS OF TEST I—PRESSEY X-O TEST

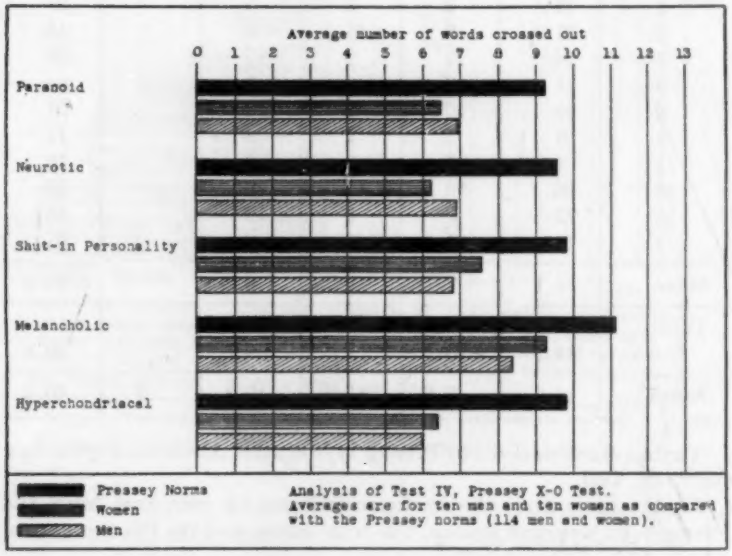

Average number of words crossed out
0 1 2 3 4 5 6 7 8 9 10 11 12 13
Paranoid
Neurotic
Shut-in Personality
Melancholic
Hyperchondriacal
Pressey Norms
Women
Men
Analysis of Test IV, Pressey X-O Test.
Averages are for ten men and ten women as compared with the Pressey norms (114 men and women).

CHART 3. ANALYSIS OF TEST IV—PRESSEY X-O TEST

TABLE 3

Analysis of Test IV—Pressey X-O Test

CASE NUMBER	PARANOID	NEUROTIC	SHUT-IN	MELANCHOLIC	HYPERCHON	TOTAL
			Women			
4	10	6	8	14	7	45
7	1	2	1	3	0	7
8	6	4	15	12	12	49
10	2	3	2	4	0	11
12	7	6	8	6	10	37
13	7	12	11	12	13	55
15	13	8	11	13	8	53
16	6	4	7	7	3	27
17	8	3	7	16	8	42
20	4	14	7	7	3	35
Mean.....	6.4	6.2	7.7	9.4	6.4	36.1
			Men			
1	16	9	12	15	10	62
2	3	0	3	2	1	9
3	9	7	7	11	3	37
5	3	8	6	3	8	28
6	1	5	0	1	2	9
9	4	7	2	8	8	29
11	4	5	6	11	7	33
14	10	13	15	13	3	54
18	10	6	11	7	2	36
19	8	8	5	13	16	50
Mean....	6.8	6.8	6.7	8.4	6.0	34.7
Total Mean...	6.6	6.5	7.2	8.9	6.2	35.4
Norm.....	9.3	9.6	9.7	11.1	9.8	49.5

Tentative norms given by Pressey in Journal of Abnormal Psychology for April, 1921.

Each column gives the words crossed out by each case with the averages for men and women, the total cases, and the Pressey norms.

TABLE 4

Deviation on Brotemarkle Comparison Test

CASE NUMBER	I	II	III	IV	V	VI	VII	TOTAL
Women								
4	2	2	4	8	2	8	2	28
7	2	4	4	2	0	4	12	28
8	0	2	6	10	2	6	4	30
10	2	6	4	2	2	8	14	38
12	6	4	4	6	6	8	6	40
13	2	10	2	0	0	0	4	18
15	4	6	0	4	0	4	0	18
16	4	6	2	2	6	2	0	22
17	2	2	4	10	2	6	12	38
20	0	4	0	4	0	4	2	14
Mean.....	2.4	4.6	3.0	4.8	2.0	5.0	5.6	27.4
Men								
1	4	2	2	8	4	4	4	28
2	0	4	8	6	0	6	6	30
3	4	4	8	2	2	6	4	30
5	2	2	0	4	0	6	4	18
6	0	4	2	8	0	6	0	20
9	0	0	2	2	2	4	8	18
11	4	2	2	4	4	2	4	22
14	0	4	6	4	6	8	2	30
18	4	4	0	14	2	4	4	32
19	2	6	2	0	4	2	2	18
Mean.....	2.0	3.2	3.2	5.2	2.4	4.8	3.8	24.6
Total Mean...	2.2	3.9	3.1	5.0	2.2	4.9	4.7	26.0
Norm.....	3.0	4.1	3.1	4.4	3.4	4.5	3.3	24.0

I Basic Moral Principle
II Reactive Profile
III Emotional Response
IV Expressive-Regressive Attitude
V Relational Adjustment
VI Underlying Feeling-Emotive
VII Resultant Orientation

TABLE 5

Five test scores

CASE	A*	I	Q	T	B	TOTAL
			Women			
4	235	38	108	57	28	466
7	159	39	126	78	28	430
8	206	43	114	75	30	468
10	87	41	114	73	38	353
12	185	58	115	66	40	464
13	209	41	114	61	18	443
15	227	61	118	72	18	496
16	155	41	109	74	22	401
17	216	37	111	81	38	483
20	196	49	118	60	14	437
Mean.....	187.5	44.8	114.7	69.7	27.4	454.1
			Men			
1	191	49	117	60	28	445
2	142	54	107	73	30	406
3	204	40	113	72	30	469
5	154	39	115	76	18	402
6	175	50	111	84	20	440
9	131	45	110	62	18	366
11	134	42	120	73	22	391
14	293	40	121	60	30	544
18	193	48	122	73	32	468
19	227	47	112	68	18	472
Mean.....	184.4	45.4	114.8	70.1	24.6	440.3
Total Mean...	185.9	45.1	114.7	69.9	26.0	447.2
Median...	230	47.2	109		24.0	

*A—"Total Affectivity" score on Pressey X-O Test.
I—"Total Idiosyncrasy" score on Pressey X-O Test.
Q—IQ as determined by the Otis Scale.
T—Total score on Downey Will-Temperament Test.
B—Total "deviation" on Brotemarkle Comparison Test.
Medians as given by the authors of the tests.

why is it that no significant correlations are found between her total scores and those of the other tests? (See table 6.) The same question might be asked concerning the other test.

The lack of correlation between the five sets of test scores is rather striking. One of two conclusions is possible: either the tests are not all that their authors claim them to be, or else the method of checking their validity was at fault.

It should be said that the authors of all three tests admit that the total scores are not the important data to be secured by

TABLE 6

Correlations of five test scores

		P.E.
T B*	.173	.15
A I	.008	.15
A Q	.135	.14
A T	−.341	.13
A B	−.031	.15
Q I	.007	.15
T I	−.051	.15
B I	.088	.14
Q T	.008	.15
Q B	.019	.15

*A—"Total Affectivity" score on Pressey X-O Test.
I—"Total Idiosyncrasy" score on Pressey X-O Test.
Q—IQ as determined by Otis Scale.
T—Total score on Downey Will-Temperament Test.
B—Total "deviation" on Brotemarkle Comparison Test.

their tests. Downey urges the study of the Will-Temperament profile and points out the fact that the total score may be misleading. Pressey, although he gives tentative norms for total scores, points to the much greater value of analysis of each individual test.[3] Brotemarkle admits that a total deviation score on his test is practically valueless unless the deviation on each of the seven series of the test is known.[4]

[3] *Journal of Abnormal Psychology*, April, 1921.
[4] *Journal of Applied Psychology*, February, 1928.

If the method used is valid, we should feel justified in making a rather harsh criticism of the tests. Before doing this, however, it seems only fair to give the suggestions of the authors a fair trial and actually use the analytical method of comparison for each case. With this thought in mind, the foregoing results were taken to the author nearest at hand, namely, Prof. R. A. Brotemarkle of the University of Pennsylvania. He accepted the challenge for the group, and suggested that the case-study method be applied to the same data in order to see whether it too failed to disclose significant results. The remainder of this study is the result of carrying out this suggestion.

II. THE CASE-STUDY APPROACH

Procedure

At the suggestion of Prof. Brotemarkle the data for each individual student was set out in the form of profiles and scores, making a case chart. Using these charts as a sole source of information he analyzed each case. To check the validity of this clinical analysis two other statements were secured concerning each case from men who had known the students for a period of one school year. Analysis I was from an instructor associated with the students in the classroom and in activities. Analysis II was by the dean of the school, whose contacts with the students had been in his official capacity.

In order to further check the validity and employability of the tests the authors of the other two tests were extended the invitation to make an analysis on the basis of the case charts.

Professor Downey replied, "I do not like to risk a chance on personality reading on the basis of the Will-Temperament Profile taken by itself."[5]

Professor Pressey[6] suggested that the authors take the matter up with Prof. O. R. Chambers who, he felt, would be interested in attempting such an analysis. The latter, however, did not respond to the invitation.[7]

[5] Letter of July 26, 1928.
[6] Letter of August 1, 1928.
[7] To date November 5, 1928.

Results

KEY TO CASE CHARTS

PRESSEY X-O TEST

Analysis of Test I

D—Disgust
F—Fear
S—Sex
S—Suspicion
J—Jokers

Analysis of Test IV

P—Paranoid
N—Neurotic
S—Shut-in
M—Melancholic
H—Hyperchondriacal

DOWNEY WILL-TEMPERAMENT PROFILE

A—Speed of Movement
B—Freedom from Load
C—Flexibility
D—Speed of Decision
E—Motor Impulsion
F—Reaction to Contradiction
G—Resistance to Opposition
H—Finality of Judgment
I—Motor Inhibition
J—Interest in Detail
K—Co-ordination of Impulses
L—Volitional Perseveration

BROTEMARKLE COMPARISON TEST

I Basic Moral Principle
II Reactive Profile
III Emotional Response
IV Expressive-Regressive Attitude
V Relational Adjustment
VI Underlying Feeling-Emotive
VII Resultant Orientation

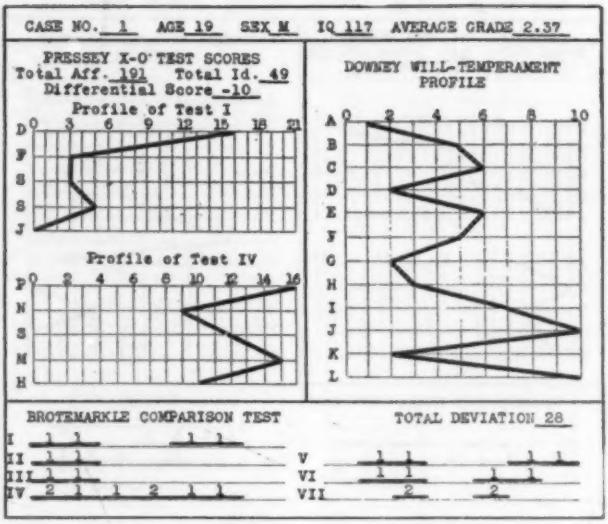
Observational Analysis I. By an instructor associated with students for one school year in classroom and in activities.

Observational Analysis II. By dean of school associated with students in that capacity for one school year.

Clinical Analysis. By a psychological clinician, R. A. Brotemarkle, accepting the challenge for the authors of the three tests. His analysis was solely on basis of data given on the case charts.

Student's Reaction to Clinical Analysis. Obtained by personal interview with each student.

CASE 1


CASE NO. 1 AGE 19 SEX M IQ 117 AVERAGE GRADE 2.37
PRESSEY X-O TEST SCORES
Total Aff. 191 Total Id. 49
Differential Score -10
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 28

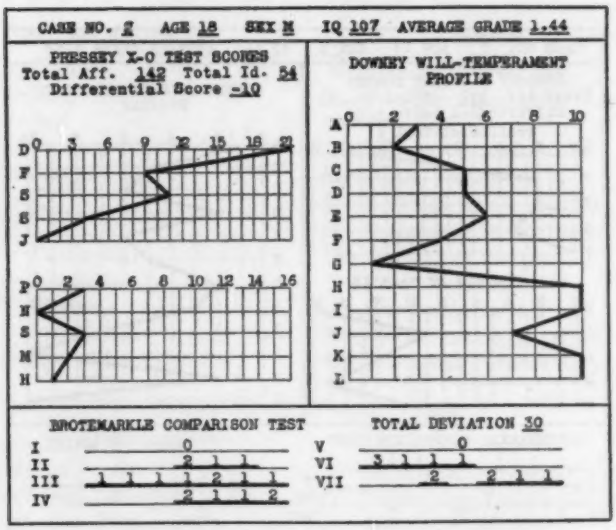
Observational Analysis I. An exceptional student, capable of doing an excellent grade of work. Reactions are slow. Probably a primary introvert. Violent temper although usually controlled; expressed by sulkiness. Well liked and very dependable.

Observational Analysis II. A good student. Intelligent. Very sensitive. A good worker but slow in movement. A good reasoner. Dependable.

Clinical Analysis. Poor Expressive-Regressive Attitude. Temperament factors high. Extreme attitude of disgust. Easily depressed and over-suspicious. Not aggressive. Apt to be personal in remarks—very biting on occasions but on others will shut up like a clam.

Student's Reaction to Clinical Analysis. A very good line-up of me. Have always been suspicious. Many people fill me with disgust by their manners, attitudes, etc.

CASE 2



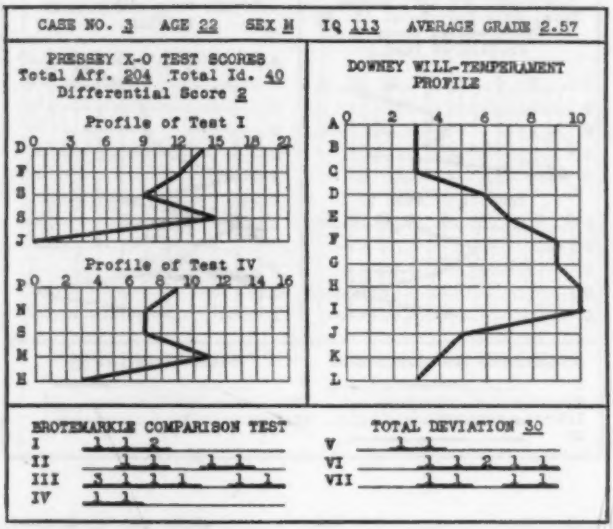
Observational Analysis I. Here is a changeable and explosive type. Never can be sure what he is going to do and are not surprised if he goes off the handle. Does average work, but at times shows ability to do high grade work. Would suspect there is some serious problem responsible for his general mal-adjustment, for he seems to have qualifications of good personality.

Observational Analysis II. A poor student. Slightly dull—may be due to habits. Somewhat sensitive. Poor in dependability. Slow in movement.

Clinical Analysis. Good emotional balance. Exaggerated disgust. Activity low. Not in the least aggressive. Excellent control. Emotions not extreme. Can handle himself in any situation. A generally good type.

Student's Reaction to Clinical Analysis. No particular reaction. Accepted analysis.

CASE 3


CASE NO. 3 AGE 22 SEX M IQ 113 AVERAGE GRADE 2.57
PRESSEY X-O TEST SCORES
Total Aff. 204 Total Id. 40
Differential Score 2
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 30

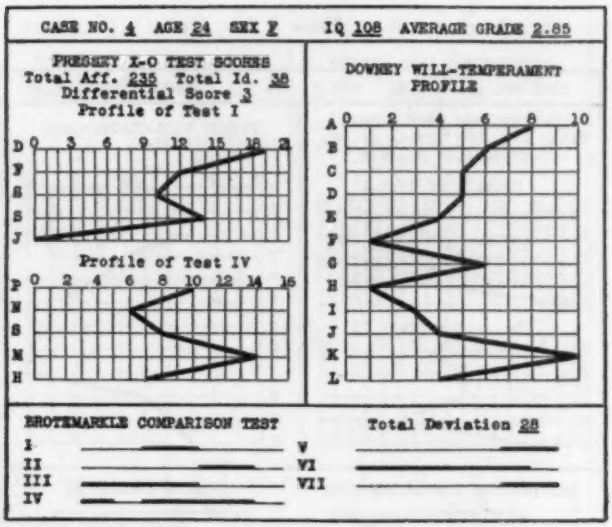
Observational Analysis I. Seems to be a normal individual. Good intelligence and does very good work in studies.

Observational Analysis II. A good student. Intelligent. Not very emotional. Dependable so far as I know.

Clinical Analysis. Well balanced. Emotional factors are high. Slight tendencies towards melancholic and self-conscious reactions. Low in activity. Good adjustment and well-balanced. Not overtly active in adjustments which are largely internal.

Student's Reaction to Clinical Analysis. Student felt it was a fair and helpful analysis. Admits melancholic tendency which has been largely overcome.

CASE 4



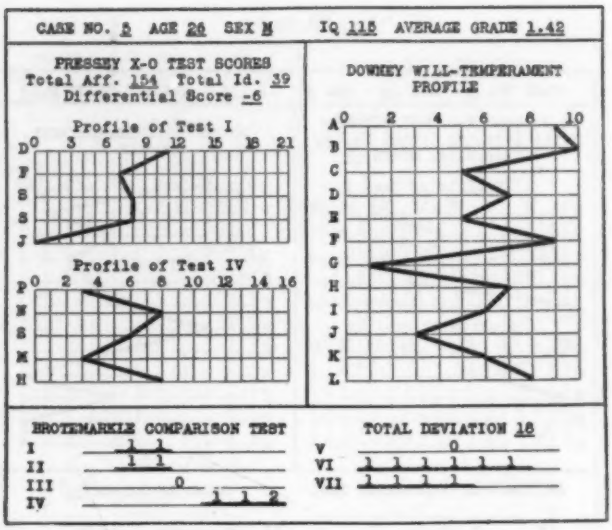
Observational Analysis I. This young lady has a very pleasing personality. Fits into the group well and is well liked. She is average as to mental activity. Seems to have plenty of assurance and purposive control. She has a cheerful disposition. It seems that she has times of depression, but her general air is that of one who takes life as it comes and does not worry.

Observational Analysis II. A good student. Quick of movement and speech. Fair in intelligence. Emotional. Fairly trust-worthy. Unscientific in attitude. Credible. Inclined sometimes to depend on intuition and not to use reason.

Clinical Analysis. Aggression very low. An overload of disgust with corresponding overload of hate. Emphatic melancholic type. Underlying feeling emotive low.

Student's Reaction to Clinical Analysis. Student surprised at accuracy of analysis. Admits depression which she is trying to hide. Attributes disgust and hate to home conditions.

CASE 5


CASE NO. 5 AGE 26 SEX M IQ 115 AVERAGE GRADE 1.42
PRESSEY X-O TEST SCORES
Total Aff. 154 Total Id. 39
Differential Score -6
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 16

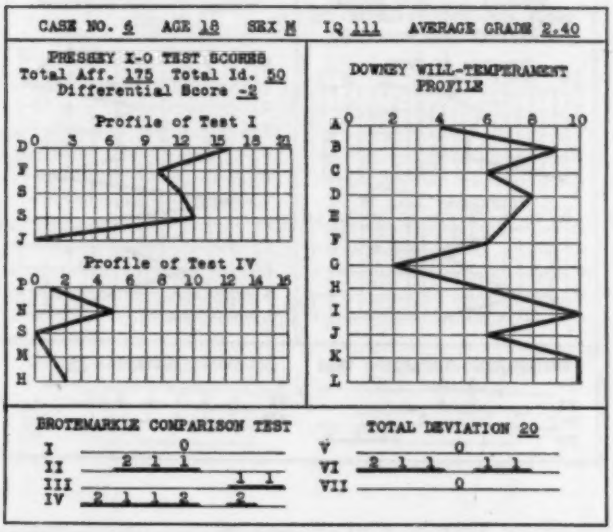
Observational Analysis I. An extremely aggressive individual who likes to be in the front ranks in any situation. Dogmatic. With complete assurance. Fits into the group fairly well, although, as we should expect from this type, has frequent minor clashes. Seems capable of doing a high grade of work, but apparently is too much interested in other things to really take his work seriously. A primary introvert.

Observational Analysis II. Intelligence fair. Slightly below average as a student. Perhaps does not concentrate his attention on his studies. Good initiative. Fair in co-operation. Excellent in expression. Slightly egotistic. Good in leadership.

Clinical Analysis. High strung, active individual. Probably involves a sickly background. The whole balance is nearly normal.

Student's Reaction to Clinical Analysis. Student tells of sickly background which surrounded life up to late in high school. Cannot understand how this was discovered, since he now enjoys good health.

CASE 6


CASE NO. 6 AGE 18 SEX M IQ 111 AVERAGE GRADE 2.40
PRESSEY X-O TEST SCORES
Total Aff. 175 Total Id. 50
Differential Score -2
Profile of Test I
D F S S J
Profile of Test IV
P N S M H
DOWNEY WILL-TEMPERAMENT PROFILE
A B C D E F G H I J K L
BROTEMARKLE COMPARISON TEST
I 0
II 2 1 1
III 1 1
IV 2 1 1 2 2
TOTAL DEVIATION 20
V 0
VI 2 1 1 1 1
VII 0

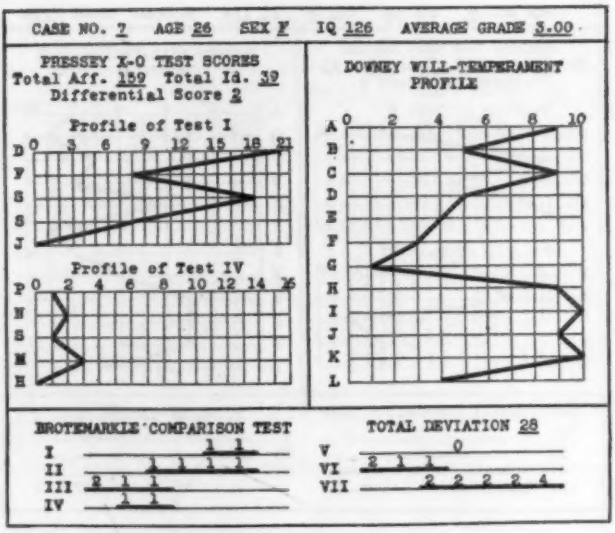
Observational Analysis I. This young man is attractive and pleasant in his manner. Apparently he has no problems nor does he cause any. Is not aggressive, but on the whole modest and retiring. All in all a normal, stable type.

Observational Analysis II. Good intelligence. A good student. Co-operative. Fairly good in leadership. Not very emotional. Dependable.

Clinical Analysis. Deviates on Expressive-Regressive Attitude, which indicates shyness. Slightly neurotic. On the whole a balanced temperament.

Student's Reaction to Clinical Analysis. Satisfied with the analysis, but objects to being labeled as shy.

CASE 7


CASE NO. 7 AGE 26 SEX F IQ 126 AVERAGE GRADE 3.00
PRESSEY X-O TEST SCORES
Total Aff. 159 Total Id. 39
Differential Score 2
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 28

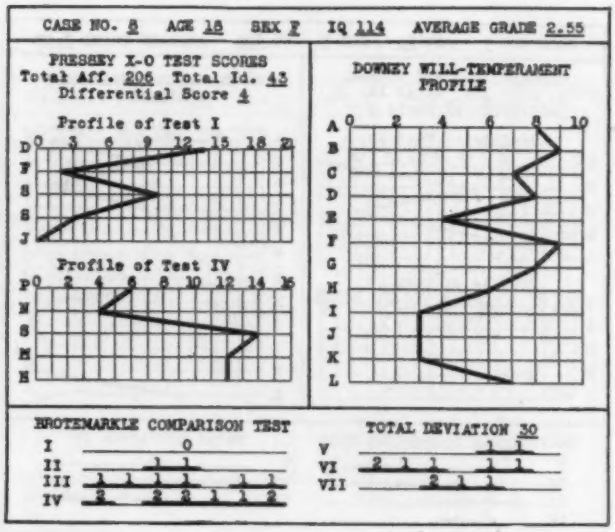
Observational Analysis I. This student has a splendid personality and high intelligence. Is well liked and one who would be prized as an acquaintance. Shows self-confidence and assurance, but does not push herself forward offensively. Is very ambitious in the right way and shows great capability in many directions.

Observational Analysis II. An excellent student. Intelligent. Good initiative. Co-operative. Quick in movement. Emotional. Very dependable. Good in leadership.

Clinical Analysis. Deviates on orientation, on sex, and on disgust. Fair temperament, smooth and balanced. Having a terrific struggle in her inner life. Trying to appear indifferent to sex which is her prime interest in life. Anxious to be ambitious. Fights problem of indifference.

Student's Reaction to Clinical Analysis Admits a sex problem (disappointed in love) but thought that she had settled it many years ago by throwing whole life into career. Surprised to find it analyzed.

Case 8


CASE NO. 8 AGE 18 SEX F IQ 114 AVERAGE GRADE 2.55
PRESSEY X-O TEST SCORES
Total Aff. 206 Total Id. 43
Differential Score 4
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 30

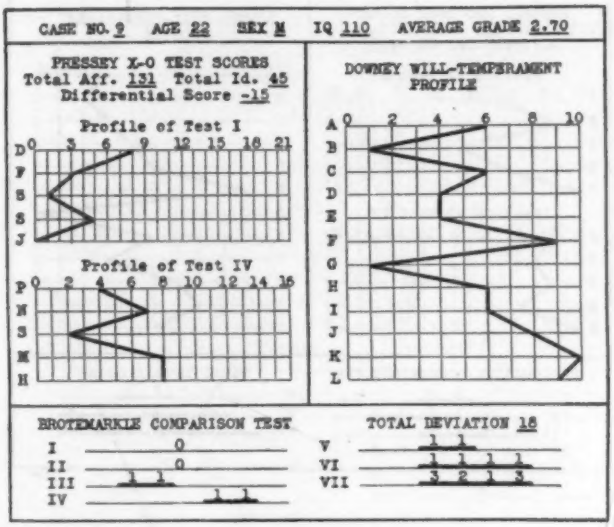
Observational Analysis I. Very active type. Average intelligence but not much disposition to use it. Has a serious lack in regard to propriety; being apt to get into all kinds of trouble. Thinks after acting wrongly, rather than before.

Observational Analysis II. Intelligent. Excellent ability which is not always used. Initiative good. Inclined to resent direction. Flighty. Flippant. Inclined to attempt to deceive. Emotional. Affectionate, lighthearted, gay, happy.

Clinical Analysis. Extreme deviation on Expressive-Regressive Attitude. Mal-adjustment is extreme. Slight emotional tinge as background. Shut-in type. Self-conscious, depressed, hyperchondriacal background. A primary introvert. Can rationalize, but cannot function because of emotional background. Lacks will power. Going to have trouble in most situations in life. If there is time to rationalize, will pass in most situations, but under pressure will fail.

Student's Reaction to Clinical Analysis. Admits her greatest weakness in inability to look ahead and see consequences of actions. Many troubles caused by it.

CASE 9


CASE NO. 9 AGE 22 SEX M IQ 110 AVERAGE GRADE 2.70
PRESSEY X-O TEST SCORES
Total Aff. 131 Total Id. 45
Differential Score -15
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 16

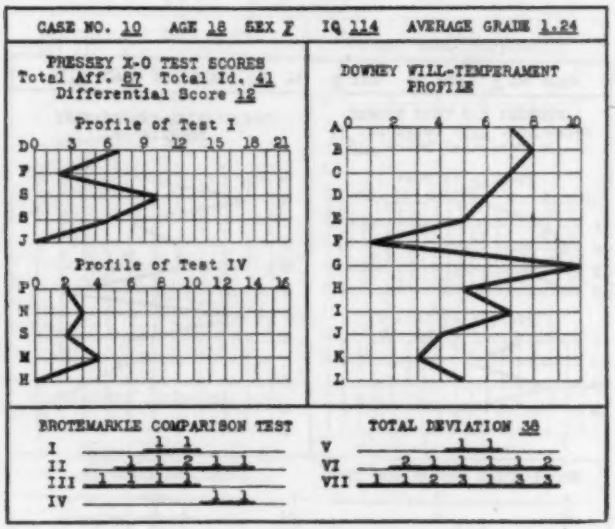
Observational Analysis I. This young man has average ability coupled with a serious purpose, enabling him to turn out high grade work. Is pleasant and interesting in manner, although somewhat retiring. The dependable sort of individual.

Observational Analysis II. A good student. Intelligent. Co-operative. Good in leadership. Not very emotional. Very dependable.

Clinical Analysis. Somewhat poor orientation. A normally balanced temperament, with perhaps slight tendency to depression.

Student's Reaction to Clinical Analysis. Feeling of instability because not settled on vocation, although Junior in college. Depression due to home background.

Case 10


CASE NO. 10 AGE 18 SEX F IQ 114 AVERAGE GRADE 1.24
PRESSEY X-O TEST SCORES
Total Aff. 87 Total Id. 41
Differential Score 12
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 38

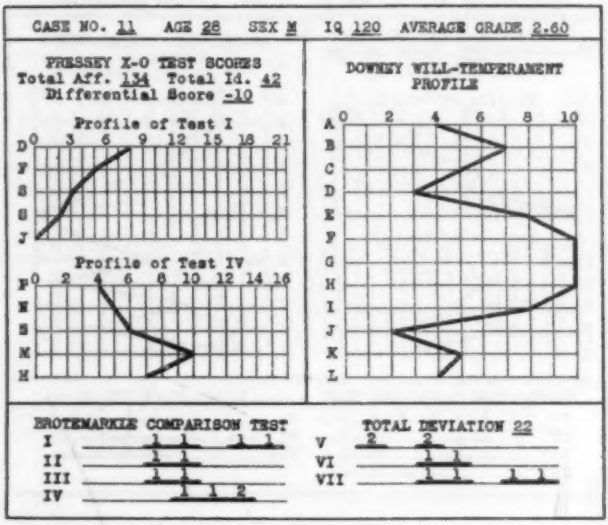
Observational Analysis I. Here is a case of an individual as changeable as the winds themselves. Seems to lack any decisive control. Is undoubtedly capable of doing fair work, but is apparently too lazy to make the effort. Self-conscious in a high degree. Very apt to lose control. The weepy type.

Observational Analysis II. Slow of movement. Not very emotional. Intelligence fair. Slightly below average in reasoning. Fairly easily influenced for the right. Requires coaxing. Fairly trustworthy.

Clinical Analysis. Balanced temperament. Slight emphasis on sex; correlating with emphasis on disgust and self-consciousness. Fairly smooth emotional curve. Am suspicious that she has a problem which might be the balance of sex to life, due to adolescence. Lazy. Slow development. Probably this case represents a slow development of adolescence. Upon completion of development, may be brilliant success.

Student's Reaction to Clinical Analysis. Admits self-consciousness. Objects to being called lazy.

CASE 11


CASE NO. 11 AGE 28 SEX M IQ 120 AVERAGE GRADE 2.60
PRESSEY X-O TEST SCORES
Total Aff. 134 Total Id. 42
Differential Score -10
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
EROTEMARKIE COMPARISON TEST
TOTAL DEVIATION 22

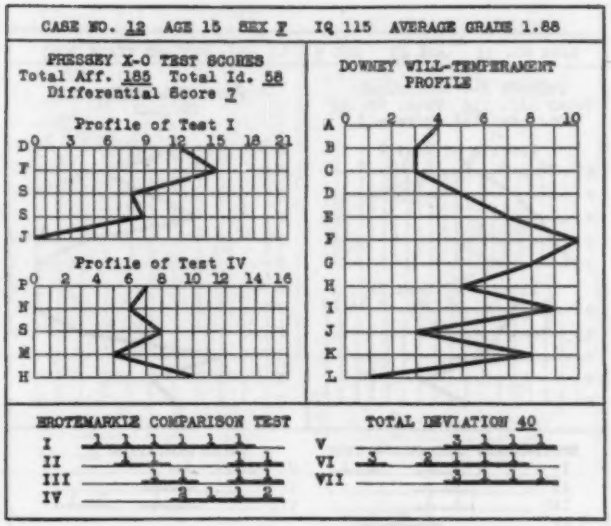
Observational Analysis I. This person is deliberate but purposeful and conscientious. He has a fine intellect and plenty of ambition to use it. Does excellent class-room work. Is well liked by the group.

Observational Analysis II. An excellent student. Good intelligence. Co-operative. Very dependable. Not very emotional.

Clinical Analysis. Good balance in emotional response, with no specific defect. Normal in emotional spread. Fair activity. Emphatic aggressive background.

Student's Reaction to Clinical Analysis. Student questions aggressiveness at first, but upon reflection agrees that the analysis is a fair one.

CASE 12


CASE NO. 12 AGE 15 SEX F IQ 115 AVERAGE GRADE 1.88
PRESSEY X-O TEST SCORES
Total Aff. 185 Total Id. 58
Differential Score 7
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 40

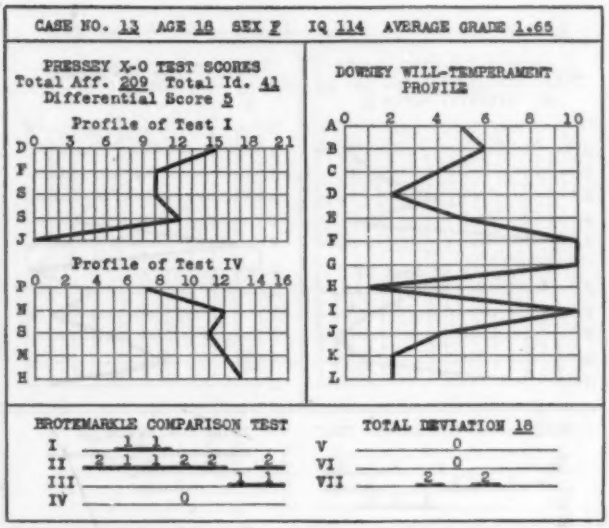
Observational Analysis I. This case is a hard one to judge, since she is still immature. She has good intelligence and does excellent work in classroom. In general a rather easy-going type.

Observational Analysis II. A fair student. Fair in intelligence. Buoyant in spirit. Co-operative. Easily lead. Fair in dependability.

Clinical Analysis. Deviation in Expressive-Regressive Attitude, Basic Moral Judgment, and Orientation. Slight mal-adjustment in emotional response. Basically emotional and intellectual as opposed to physiological. Activity low. Typical case of slight mal-adjustment in unbalanced adolescence. Tries to rationalize everything.

Student's Reaction to Clinical Analysis. Has trouble getting along with folks, expressed internally. Some doubting.

CASE 13



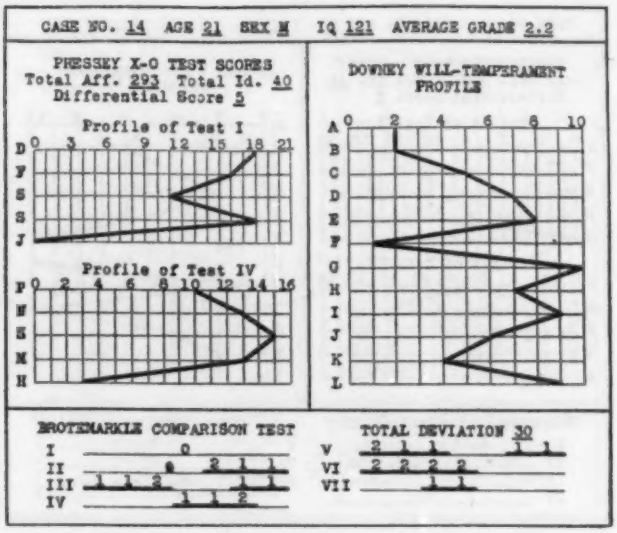
Observational Analysis I. Has average ability. Shows occasional outbursts of enthusiasm or depression. Plenty of assurance which is apt to show itself in stubbornness. Not apt to be popular.

Observational Analysis II. A good student. Intelligent. Not very emotional. Fairly easily influenced. Fair in dependability, unless influenced by others.

Clinical Analysis. Reaction profile reveals a physiological background. Unstable emotionally; but the temperament defect is physiological. The disturbance may be internal, and any trouble will be adjustment within herself rather than with the group. Conduct in group usually satisfactory.

Student's Reaction to Clinical Analysis. Sickly background and general weak physical condition which cause worry and depression.

CASE 14


CASE NO. 14 AGE 21 SEX M IQ 121 AVERAGE GRADE 2.2
PRESSEY X-O TEST SCORES
Total Aff. 293 Total Id. 40
Differential Score 5
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 30

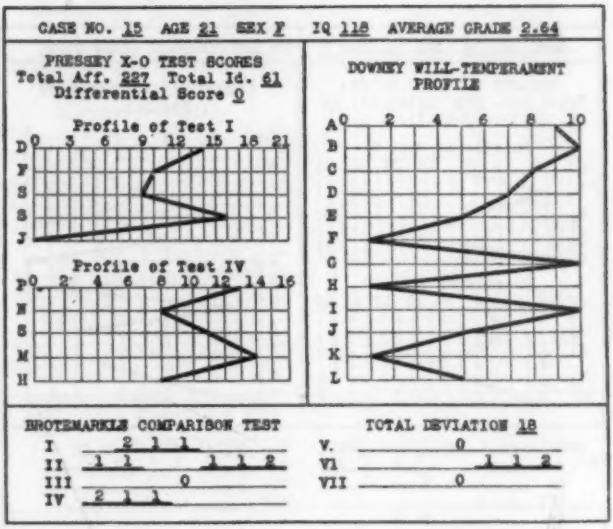
Observational Analysis I. A strong aggressive type, apt to border on the dogmatic, however. Has good ability and can do an excellent piece of work when he so desires. Is not especially popular in the group, which is probably due to his self-satisfied air.

Observational Analysis II. A good student. Fair in intelligence. Not very emotional. Fair in reasoning ability. Trustworthy.

Clinical Analysis. Deviation in emotional response. High-strung, neurotic, and self-conscious. Temperament unbalanced. Introvert type. Slight activity with high aggressive control. Excitable. Prone to be in all sorts of mal-adjustment.

Student's Reaction to Clinical Analysis. This student was chagrined at the uncanny way in which the secrets of his life, which he was attempting to hide and overcome, were brought to light.

CASE 15



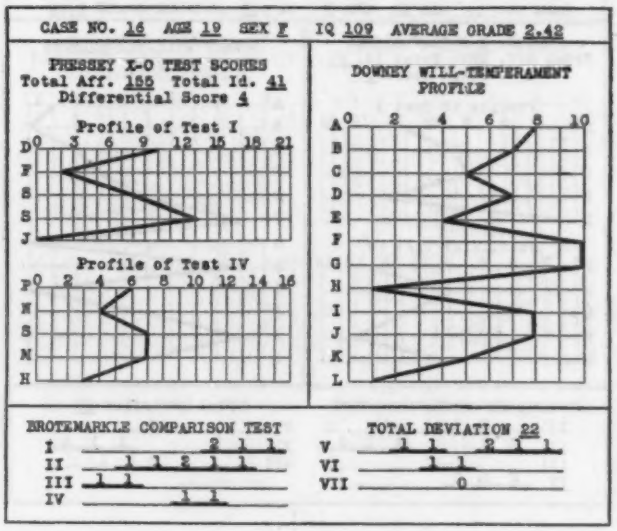
Observational Analysis I. A very active type who likes to be a leader, but often shows poor judgment in doing it. Is apt to flare up on slight provocation. At times indifferent, but may become very enthusiastic over certain things on occasions. Is moody and changeable.

Observational Analysis II. A good student. Intelligent. Reasoning not always sound. Emotional. Sensitive. Co-operative. Good in dependability.

Clinical Analysis. Very good balance and excellent emotional response. Slight fluctuation within normal range, with fluctuations in control. Borders on type which is apt to change in a second—one thing now and something else next moment. Schizophrenic type.

Student's Reacton to Clinical Analysis. Admits fluctuation, but objects to being called extreme changeable type.

CASE 16


CASE NO. 16 AGE 19 SEX F IQ 109 AVERAGE GRADE 2.42
PRESSEY X-O TEST SCORES
Total Aff. 155 Total Id. 41
Differential Score 4
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 22

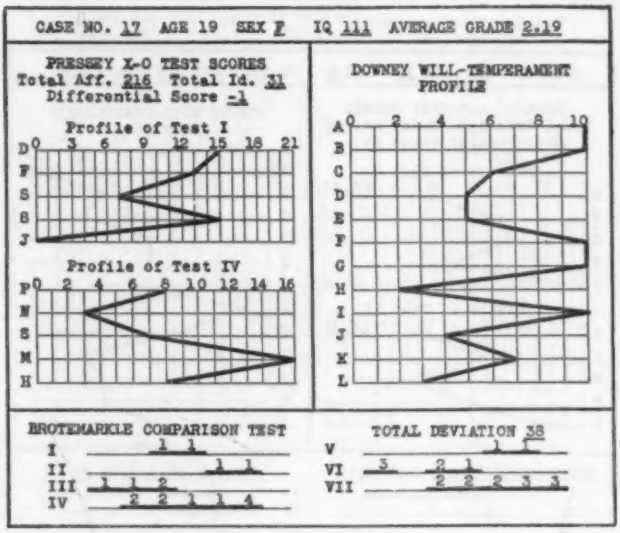
Observational Analysis I. Average intelligence with not too much ambition. A likeable, easy-going sort of person. Has a pleasing way and would always fit into a group.

Observational Analysis II. About an average student or slightly below. The same in intelligence. Not very emotional. Somewhat retiring in disposition. Slightly below average in dependability.

Clinical Analysis. The same type as Case 14. More self-conscious but less temperamental. The fluctuation is rather mental than emotional.

Student's Reaction to Clinical Analysis. Analysis points out some things correctly, principally high-strung type and self-consciousness.

CASE 17


CASE NO. 17 AGE 19 SEX F IQ 111 AVERAGE GRADE 2.19
PRESSEY X-O TEST SCORES
Total Aff. 216 Total Id. 31
Differential Score -1
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 38

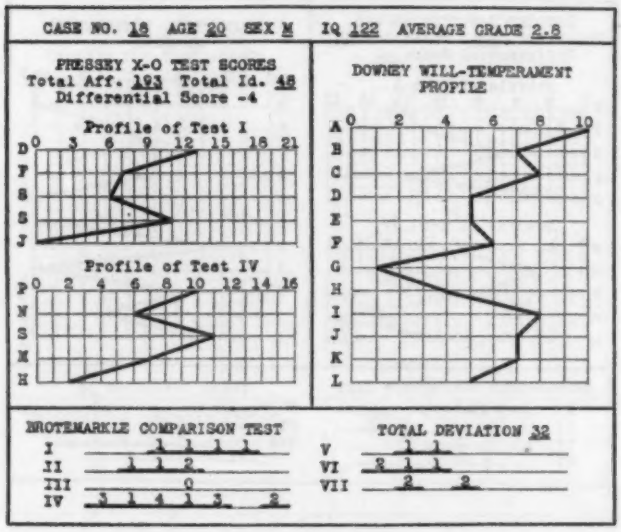
Observational Analysis I. Average intelligence, but an aggressive manner and self-confidence which is deceiving. One would expect rather high IQ. Capable but lacks ambition. High-strung, nervous type. Apt to be critical and dissatisfied.

Observational Analysis II. Intelligent, excellent in control ability, though not always used. Resents direction and especially dictation. Emotional. Behavior directed by intelligence. Inclined to deceive. Poor in dependability.

Clinical Analysis. Fluctuating type, comparable to Cases 12 and 14. Apt to be lazy Feeling background is off.

Student's Reaction to Clinical Analysis. Self-conscious in extreme degree—due to internal physiological condition. Prone to worry and lose control emotionally.

CASE 18


CASE NO. 18 AGE 20 SEX M IQ 122 AVERAGE GRADE 2.8
PRESSEY X-O TEST SCORES
Total Aff. 193 Total Id. 48
Differential Score -4
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 32

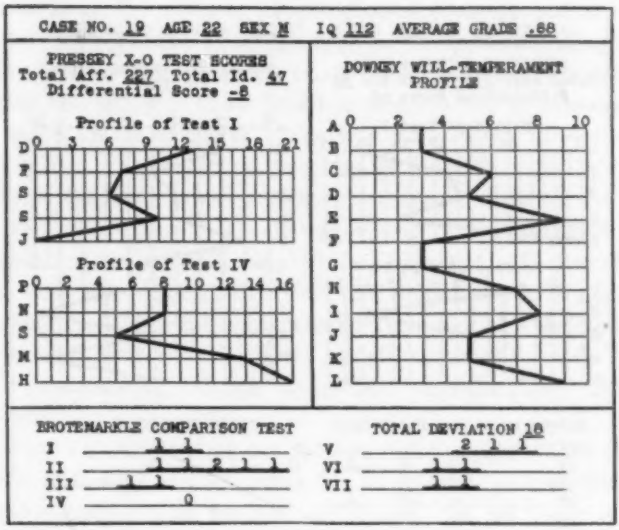
Observational Analysis I. Strong intelligence but apparently weak control. Apt to flare up. Seems to be lacking in judgment. Does not seem to fit into group very well.

Observational Analysis II. A good student. Fair in intelligence. Emotional. Slightly sensitive. Co-operative.

Clinical Analysis. Very low on Expressive-Regressive Attitude. Does not know how to react to anybody. Shut-in self-conscious type. Apt to "get in dutch." Lacks knowledge of proper response in given situation.

Student's Reaction to Clinical Analysis. No interview.

CASE 19


CASE NO. 19 AGE 22 SEX M IQ 112 AVERAGE GRADE .68
PRESSEY X-O TEST SCORES
Total Aff. 227 Total Id. 47
Differential Score -8
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 16

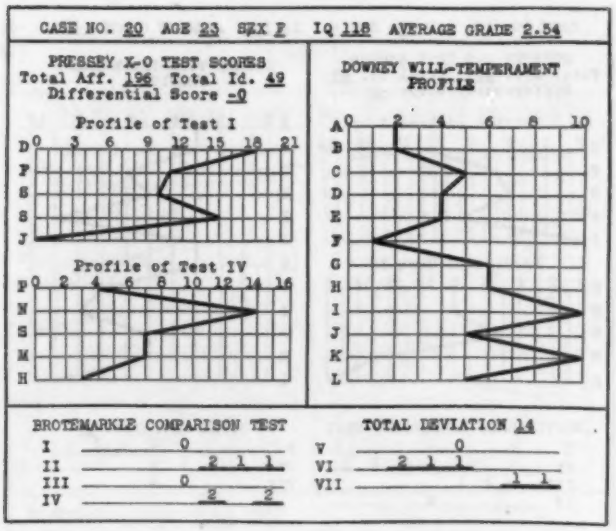
Observational Analysis I. This case is a hard one to describe, since the individual is very changeable. Is easily upset and prone to be offensively aggressive when will is crossed. On the other hand, can be a perfect gentleman in all respects if he chooses. Apparently some serious defect in control.

Observational Analysis II. Slow in movement. A poor student. Does not always study. Habits not good. Sulky. Desirous of having his own way. Poor in dependability. Fair intelligence.

Clinical Analysis. Good emotional response. Evenly balanced will-temperament. Fair balance in emotion and temperament spreads. Melancholic and hyperchondriacal background. Primary introvert whose conduct is highly over-emphasized. Apt to force himself. He is smooth enough to hold his balance in most situations.

Student's Reaction to Clinical Analysis. Student flattered by analysis. No particular reaction.

CASE 20


CASE NO. 20 AGE 23 SEX F IQ 118 AVERAGE GRADE 2.54
PRESSEY X-O TEST SCORES
Total Aff. 196 Total Id. 49
Differential Score -0
Profile of Test I
Profile of Test IV
DOWNEY WILL-TEMPERAMENT PROFILE
BROTEMARKLE COMPARISON TEST
TOTAL DEVIATION 14

Observational Analysis I. This individual has good intelligence and a pleasing personality. Gets along well with the group. Takes life seriously and consequently worries a great deal. Very self-conscious and lacks aggression. Capable and conscientious, however, and with good underlying purpose.

Observational Analysis II. A good student save for physical handicap, viz., eyesight. Good intelligence. Fair in reasoning. Good in leadership ability. Emotional. Rather sensitive.

Clinical Analysis. Normal fluctuating balance. Slightly self-conscious. Low activity with high control. Neurotic type. Can be very nice, but, on the other hand, is capable of being very nasty if control breaks down.

Student's Reaction to Clinical Analysis. Feels self-consciousness is one of her weaknesses. Very nervous, especially when tired. General health not up to par.

Discussion

The reader will find it very interesting to draw his own conclusions as to the validity of the clinical analysis and the correlation between the three analyses. It appears that on the whole there is just as much relationship between the clinical analysis and either of those based on observation as between the two based on observation. There is no striking difference of opinion in any case except, perhaps, in case 2.

The writer interviewed personally each case after all the data had been assembled.[8] The purpose was to check the clinical analysis by the student's own opinion of himself. In practically every case the student felt that the analysis was a fair and accurate one. Only a few objections were raised by the students interviewed, and these will be noted under the head, "Student's Reaction to Clinical Analysis," placed under "Clinical Analysis."

One case stands out as concrete evidence of the value of the clinical analysis, namely, case 8. It will be noticed that all three analyses are in perfect agreement. It might be added that this young lady was in serious trouble several times during the year. If the results of the clinical analysis had been in the hands of the administration at the opening of the year, much of the mal-adjustment might have been eliminated.

Perhaps the most significant fact revealed by the case studies was the fact that the clinical analysis gave important information not even hinted at by either of the observational analyses. This was particularly true in four cases, nos. 1, 5, 7 and 14. The clinical analysis of case 1 states that the individual has an extreme attitude of disgust and is depressed over suspicions. In the interview, this person admitted that he felt disgust or very strong dislike for no less than five of the students who are his daily associates. He also confessed that his boyhood was spent in a community where he had, sooner or later, to engage in fistic combat with every boy in the school. The analysis of case 5 indicates a sickly background. The interview confirmed

[8] Unable to interview case 18.

this statement, the young man having been sickly all through his school days. Neither observer would have ever suspected it, since the young man at the present time has an apparently strong body. Case 7 presents an interesting example of the value of the clinical analysis. The analysis states that the girl is trying to appear indifferent to sex—really the prime interest in her life. Neither of the observational estimates so much as hint at any problem in her life. Yet upon interview, this young woman admits that she was disappointed in a love affair several years ago, and is throwing her energy into religious work as a compensation reaction. She could not understand how the test could discover it, for she had thought that this problem in her life had been settled many years ago. Case 14 should be mentioned since the individual admitted that the analysis fitted him in every particular. He was sorry, however, that the writer had found out about his real self, since he had been trying to cover it up.

CONCLUSIONS

Certain conclusions are indicated as follows:

1. The statistical method of comparison of averages and total scores is not ample as a check on the validity of the tests examined.

2. In the hands of one dependent upon statistical test measurements and unskilled in clinical analysis the tests examined fall far short of their intended use.

3. In the hands of an experienced clinician, the scores from the tests examined serve as a basis for reasonably accurate and valuable analysis.

4. Under certain conditions it would seem feasible for one skilled only in the administration and scoring of these tests to compile the data on a single sheet, which may be sent to an experienced clinician for analysis.

WHAT IS MEASURED BY THE IOWA "APTITUDE" TESTS?

T. A. LANGLIE

Wesleyan University

In an article entitled Analysis of the Iowa Placement Tests, published in the JOURNAL OF APPLIED PSYCHOLOGY (1926, vol. x), the writer concluded among other things, that

(1) Each test is dependent to a fair degree upon intelligence, particularly the aptitude tests.
(2) A particular aptitude is dependent upon training in that particular line.
(3) Aptitude (is) general intelligence plus training.

The placement tests studied were the Chemistry Aptitude, Chemistry Training, English Aptitude, English Training, Mathematics Aptitude, and Mathematics Training, and the above conclusions were derived by correlational studies, zero order and first order partials. By "Intelligence" was meant that which is measured by our so-called intelligence tests, and by "training" was meant any experiences which had resulted in a modification of the individual's habit systems, whether it be the taking of a course in chemistry, mathematics, or English, or other less formal methods of instruction, such as reading of scientific or semi-scientific journals and books, etc.

Stoddard, in an article published in the Seashore Commemorative Number of the Psychological Monographs (vol. xxxix, no. 2, 1928), criticizes the above mentioned conclusions, as being due to an error of sampling. He says,

Langlie's second statement is essentially fallacious, for it arises out of an accident of sampling. All subjects who took the aptitude tests had really studied the subject matter for which the tests were designed to measure aptitude. For example, the same students took both the Chemistry Aptitude and Chemistry Training Examinations. Since they have studied chemistry in high school, there was a correlation between chemistry aptitude and chemistry performance. But the Chemistry Aptitude Examination is designed principally for those with no training in chemistry; when given to such students it surely becomes meaningless to say that aptitude *depends upon* training.

His criticisms of conclusions (1) and (2) are based on the same foundation as the above criticism, and although the discreetness of these aptitude tests has not been shown, he implies that the tests are "specific."

If the aptitude tests do not depend upon training, students who have had no chemistry training should score as high on the average as students who have studied chemistry in high school. But such is not the case, for students who have studied chemistry before entering college score significantly higher in the aptitude test than do students who have had no previous scholastic training. The following table reveals the extent of these differences, as well as the lack of differences in "intelligence."

Means, S.D's, and S.D's of the Differences between Means, for various groups of students as measured by an intelligence test, and by the Iowa Chemistry Aptitude Examination.

1924

TEST	GROUP	N	MEAN	S.D.	$S.D._m$	Diff.	S.D. diff.	$\frac{D}{S.D.\text{ diff.}}$
Intelligence	Chem. 14	75	57.45	33.50	3.87	.05	4.43	.01
	Chem. 4	167	57.50	27.80	2.15			
Chemistry Aptitude	Chem. 14	77	74.94	7.10	.81	3.20	.92	3.48
	Chem. 4	175	78.14	5.66	.43			

Similar results were obtained with the 1925 classes, as follows:

TEST	GROUP	N	MEAN	S.D.	$S.D._m$	Diff.	S.D. diff.	$\frac{D}{S.D.\text{ diff.}}$
Intelligence	Chem. 14	63	55.40	32.40	4.08	4.50	4.71	.96
	Chem. 4	153	59.50	29.20	2.36			
Chemistry Aptitude	Chem. 14	61	61.85	14.40	1.84	7.65	2.12	3.61
	Chem. 4	168	69.50	13.55	1.05			

It is apparent from an inspection of the above comparisons that the groups do not differ in terms of "general intelligence," but do differ significantly in terms of the Chemistry Aptitude Examination scores, the differences between the means being more than three times the standard deviation of the difference, and in favor of the Chemistry 4 groups who had studied chemistry in high school.

The only conclusion that seems possible is that these aptitude tests measure "training" to some extent. The earlier analysis indicates that they also measure intelligence to some extent. Such conclusions do not detract from the value of such examinations for prognostic purposes, however. As a matter of fact, the writer has concluded elsewhere[1] that for such use measures of "training" are more valuable than measures of "aptitude."

[1] A Comparison of "Aptitude" and "Training" Tests for Prognosis, Jour. Educat. Psych., December, 1928, vol. xix, no. 9.

A STUDY OF THE PRESCHOOL CHILD'S PICTURE AND STORY BOOKS BY THE BATTERY OF TESTS METHOD

ELEANOR OLMSTEAD MILLER

Jacksonville, Illinois

In order to determine an individual's aptitude for a given vocation, the battery of test technique has been developed as described in Professor Clark L. Hull's *Aptitude Testing*. This same method may be applied successfully, however, in other fields.

In studying the preschool child's picture and story books, if we can secure a representative group of these books and have them ranked in order to form a criterion, we may then measure them in any given number of attributes, or "aptitudes," hence discovering which ones play the greatest part in determining the desirability of a given book. This knowledge built into a multiple regression equation gives us a basis upon which we can judge other and newer books as related to these which form the criterion group. In this way we use the battery of tests method, as used in aptitude testing, for a different sort of problem.

The following account of this method of studying children's books constitutes only a part of a much more comprehensive study. Many more factors enter into a child's book than can be measured objectively; such factors as story content, subject value, etc. This brief report attempts in no sense to be complete as a study of the preschool child's picture and story books.

It is to be noted that the study of the preschool child's book is made, not only from the standpoint of his own interests, important as those are, but also from the standpoint of the opinion of specialists regarding the books to be given to him. The preschool child's first books are given to him: he does not select them for himself.

To form a criterion group of books is obviously the first task

in this problem. The preschool period was divided into two parts—children below three, and those from three to five years. Sixty books were secured for the first age group. Since this group of books was to be as representative as possible, they were collected in several different ways and represent good, poor, and indifferent books. A number of books was selected from book lists of so-called desirable books for children of these ages; a number was secured by going to book stores and asking clerks to recommend books; a third number was secured which mothers had actually used with preschool children; and, finally, a small number was obtained from a group of books given one certain preschool child by relatives and friends. Sixty is about the smallest number of cases—or books—which can be used for valid statistical purposes. But in selecting the books for the second age group it became evident from preliminary rankings that it was well nigh impossible for rankers to rank anywhere near this number of books. The books for the older group are longer, more varied, and altogether more difficult to hold in mind for ranking purposes. Hence the exactness of statistical procedure is sacrificed in the second group in order to have a more valid criterion. Only twenty-five books comprise the second age group.

Ten people ranked each group of books. Five of each ten were professional men and women who were working with or particularly interested in preschool children. The other five in each ten were mothers who were particularly interested in preschool activities and had children of preschool age at the time they did the ranking.

These rankings were then summarized and converted into an array of scores on a convenient linear scale of ten units by the method devised by Professor Clark L. Hull in the University of Wisconsin Psychology Laboratory. Correlations were then found between the scores of the rankers by the Pearson-Product-Moment method with the simplified formula:

$$r = \frac{M_{AB} - M_A \cdot M_B}{\sqrt{M_{A^2} - (M_A)^2}\sqrt{M_{B^2} - (M_B)^2}}$$

The correlations averaged .40 for Group I and .36 for Group II.

In order to form a criterion for the scores of the two groups of books, the scores given the books by the various rankers were averaged. These criterion scores were then correlated with the scores from each ranker. These correlations with the criterion averaged .62 for Group I and .65 for Group II.

We now have a criterion for each group of books. The correlations found in obtaining it are low, so low as to seem discouraging when we find the efficiency of prediction from the formula

$$E = 1 - \sqrt{1 - r^2}$$

Using this formula we find that in only 22 per cent of the cases in the first group and 24 per cent of the cases in the second group can we predict from the criterion how these people would rank the book. But probably these values are as high as we may expect when we consider the intangibility of the problem. Other studies of a more definite nature growing logically out of this one should give better results.

The books in each group were measured in a number of factors. These measures were correlated with the criteria with the results given in table 1.

In computing a multiple regression equation, it is desirable that the factors we use correlate as high as possible with the criterion and as low as possible with each other. Thus we have selected certain factors which seemed to make the best showing when correlated with the criterion and then correlated them with each other. Selecting then the most important factors according to the principle mentioned above, a multiple regression equation was computed for each group, using Yule's method as worked out by Tolley and Ezekiel and used in the Psychology Laboratory at the University of Wisconsin. These equations are as follows:

$$\text{Group I: } X_1 = .025X_2 + .005X_3 - .0006X_4 + 3.82$$

In this equation we may substitute the per cent of pictures for X_2, the number stories for X_3, and the average number of

words per story for X_4 and thus obtain a value for X_1, which would compare any new book with those already ranked.

$$\text{Group II: } X_1 = .04X_2 - .008X_3 - .04X_4 + .01X_5 + 5.26$$

In this equation, the cost would be substituted for X_2, the area for X_3, the per cent of pictures for X_4, and the number of story pages for X_5.

TABLE 1

ATTRIBUTE	CORRELATION WITH CRITERIA	
	Group I	Group II
Cost	.12	.42
Length of book		.18
Width		−.30
Thickness		.06
Area	.07	−.29
Volume		−.07
Weight		−.18
Size cover picture		−.10
Number pictures	.06	−.09
Number colored pictures		−.04
Average size pictures		−.12
Percentage of pictures	.37	−.21
Total number pages	−.26	.39
Number story pages		.38
Number stories	.24	−.05
Number illustrated stories		−.11
Average number words per story	−.13	−.05

All books contain many factors which cannot be measured by any device such as that mentioned above. Here come elements of literary merit and understanding. Certain facts may be mentioned about the books which were not considered above.

1. There seems to be no relationship between the rank of a book and its presence or absence on book lists. Book lists are best when they serve only as a means of introducing books. We need in addition some objective method of considering a book.

2. Desirable characteristics of a book are timeliness in illustrations, well fitted to a child, offering opportunity for originality, and expression of individuality.

3. Another desirable trait is found to be strong rhythm or jingle quality. Thus the fundamental need for Mother Goose does and probably always will exist.

4. Undesirable qualities of books low on the scale are unattractive binding, lack of durability, and too small in size.

5. The ABC method of arrangement is an undesirable one.

6. Fairy tales had better be left until eight or nine years and then given in the original forms.

We may now bring together what facts and opinions we have gathered including correlation coefficients, means of various attributes, comments of rankers, etc., and form a score card to be used in rating children's books. Such a score card follows:

SCORE CARDS FOR BOOKS FOR PRESCHOOL CHILDREN

Decide for what age child (mental age is more reliable than chronological age) the given book is to be rated. Rate according to the following points: (Group I is for three years and under; Group II for three to five years).

I. Physical Aspects of Book.

1. Size

Group I	*Group II*
For area of about 68 sq. in., give 6 points. Subtract one point for each 5 sq. in. above or below. Minimum 0.	For area 54 sq. in., give 4 points For area 49 sq. in., give 5 points For area 44 sq. in., give 6 points Continue at 6 points to 34 sq. in. Below 34 sq. in., reduce 1 point for each 5 sq. in. reduction in area For area 59 sq. in., give 3 points For area 64 sq. in., give 2 points For area 69 sq. in., give 1 point For area 74 sq. in. and above give 0 points.

Maximum Value 6 Points

2. Weight

For weight 360 grams or less give 4 points. (Approximately 13 ounces)

For every 25 additional grams (about 1 ounce) reduce by 1 point with 0 as minimum.

Maximum Value 4 Points

3. Paper

Estimate quality—finish, durability, binding, etc.

Maximum Value 4 Points

4. Pages

Group I	*Group II*
For 53 pages give 4 points	For 135 pages give 4 points
For 57 pages give 3 points	For 145 pages give 5 points
For 62 pages give 2 points	For 155 pages give 6 points and continue at 6 points
For 67 pages give 1 point	For 115 pages give 3 points
For 73 pages give 0 points	For 105 pages give 2 points
Continue at 0	For 95 pages give 1 point
For 48 pages give 5 points	For 75 pages or less give 0 points
For 43 pages give 6 points	
Continue at 6 points to 25 pages	
Under 25 pages reduce 1 point for each 5 pages.	

Maximum Value 6 Points

Total Maximum Value 20 Points

(For Group II add 1 point if book costs $1.19 or more; 2 points if $1.50 or more).

II. Pictures

1. Size

Estimate, making decreased amount for very large or very small pictures.

Maximum Value 5 Points

2. Number

Give 3 points for 50 pictures. Increase or decrease 1 point for each ten pictures more or less and continue at maximum or minimum.

Maximum Value 5 Points

3. Percentage

Group I	*Group II*
For 54 per cent give 5 points	For 26 per cent give 5 points
For 59 per cent give 6 points	For 31 per cent give 6 points

Increase 1 point for each 5 per cent of increase with maximum at 10 points.

Decrease 1 point for each 5 per cent of decrease with minimum at 0 points.

Maximum Value 10 Points

4. Quality

Estimate quality on basis of clearness, color, line, subject, etc.

Maximum Value 10 Points

Total Maximum Value 30 Points

III. Stories

1. Length

Group I	*Group II*
Give 3 points for story of 450 words.	Give 3 points for story of 1600 words.

Increase points value if number of words is below; decrease if above.

Maximum Value 5 Points

2. Correspondence to age

Estimate as to simplicity, repetition, plot, etc.

Maximum Value 10 Points

3. Amount of rhythm

Estimate, giving Mother Goose rhymes 5 points, others accordingly.

Give 0 where no rhythm appears.

Maximum Value 5 Points

4. Subject

Estimate according to familiarity to child of age considered.

Maximum Value 10 Points

5. Vocabulary

Estimate according to age of child considered.

Maximum value 5 Points

Total Maximum Value 35 Points

IV. Author, Artist, Publisher.

Estimate, according to actual acquaintance, as far as it can be determined with children or previous success with children's books.

Total Maximum Value 5 Points

V. Construction of Books.

Estimate, according to materials, binding, etc. in connection with age for which intended.

Total Maximum Value 10 Points

Grand Total Maximum Value 100 Points

In conclusion, it may be stated that while we obviously have not discovered any great or momentous fact in regard to children's books from the method here employed, we have shown that this method—a far more accurate one than the usual methods for evaluating children's books—could be used with greater success with a more homogeneous group of books. Our problem is probably too broad and too intangible to gain much definite information. It does suggest, however, points at which the same method might be applied with promise of greater success. An investigation is already under way with Mother Goose books and promises more tangible results.

We have presented a score card which has already been used with some of the newer books with success. It should be valuable to the book reviewer or to the individual who selects children's books, in the same way as the multiple regression equation is valuable to the vocational expert in selecting individuals for a given vocation.

THE RELATIVE EFFECTS OF SIZE OF TYPE, LEADING AND CONTEXT

HOWARD T. HOVDE

University of Pennsylvania

SCOPE OF THIS INVESTIGATION

The experiment, reported in this paper, was conducted originally in an effort to determine the relative legibility of Intertype Ideal News Face Type, set under standard newspaper publishing conditions, with changes in size and leading. The results show the relative effects of size, leading and context. The experiment is divided into two parts: Part I records the direct measurement of the reading rate of 300 adults; Part II describes the readers' preferences and opinions of the legibility of the type setting.

HISTORICAL NOTES ON LEGIBILITY

A complete review of the literature on this subject would include several hundred titles. Anyone interested in more detailed references than can be included here will find two bibliographies independently compiled by Pyke (35) and by Tinker (41), with accompanying digests of the subjects. Research in this field has been undertaken by numerous educators, ophthalmologists and psychologists. The approach to the subject has been largely psychological. Banister (3), working with test type charts of ophthalmic standards, concluded that errors often attributed to physical defects of the eye are due more likely to psychological causes and that psychological tests are a particularly desirable approach to the problem.

Two favorite methods of investigation are prevalent. The first is the test for legibility of isolated letters and numerals of which the Snellen Chart of the ophthalmologist early served as an introduction to the general problem. The second is the test for legibility of continuous reading for context with which this paper is concerned.

Under the first method, the work in ophthalmology may be mentioned. Hartridge and Owen (20) discuss the merits of the Snellen and the Green test types in ophthalmic practice. Cowan (10) calls attention

to the failure of ordinary type to satisfy the psychological requirements of a test for visual acuity and makes suggestions for changes in the character and the design of letters of equal legibility. In a second paper (11) he reviews the field of test letters and geometric designs for the determination of visual acuity.

In the field of psychology, Roethlein's comprehensive study (38) of isolated letters has received wide recognition, and more recently Tinker (42–44) has been interested in the study of isolated letters, digits and mechanical signs. Crosland and Johnson (12) have studied the range of apprehension as affected by inter-letter hair spacing and by characteristics of the individual letter. In the main these tests meet the criticism of Ewing (15), in a discussion of the value of letters and characters as visual tests, who concludes that 62 per cent of normal vision is sufficient for ordinary use and that the army accepts men with 50 per cent vision of the ophthalmic standard and with 75 per cent vision for the finest rifle work. Parsons (34) rightly contends that the characteristic group of the letters are attended to in the reading process and not the individual letters. In general, the method of investigation for the legibility of isolated letters has met with criticism by those who have studied reading habits.

Under the second method, the inquiry into reading habits permits a different approach to the subject of legibility. The reading process has been studied experimentally by Javal (25), Erdmann and Dodge (14), Huey (21–23), Quantz (36), Dearborn (13), and more recently by Judd (26), Gray (18), Schmidt (40), Buswell (7–8) and others. Cattell (9), as early as 1885, produced strong evidence that reading proceeds not by successive letters but rather by whole words, phrases and sometimes short sentences as units in the perceptual process. Reading involves a series of eye movements across the line of type. The eye pauses between movements and is fixed upon a word or group of words. During these fixations comprehension takes place, as nothing is seen while the eye is in motion. At the end of the line a regressive movement is made to the first word in the next line and the eye again continues its broken sweep across the line. Experimental studies of this nature have been largely responsible for the discard of the assumption that reading takes place by the alphabetical method; namely, that the reading process proceeds by successive letters and not by word, phrase and sentence wholes.

The reading of any consecutively written material relies upon context for the derivation of its meaning. In this usage the word "context" includes that which might be amplified into the term "contextual meaning" but which is inferred by the single word. Buswell (7) has given a clear analysis of the complex character of meaning underlying the general process of thought that is included in the context, and Huey (23) has similarly interpreted this idea.

In considering the sensory content of reading material, a variety of conflicting opinions and contradictory recommendations have arisen from experimental evidence. A few of these observations may be considered.

Size and Face of Type: Pyke (35) experimented with 8 different type faces and found rather large legibility differences in one test and small and conflicting differences in the other tests. He concluded that extremely large typographical differences must be present before it is possible to discern differences in legibility. He used the tachistoscopic method of exposing the reading material and doubts whether the differences in legibility found in the laboratory experiments hold true under ordinary reading conditions. Huey (23) concluded that the size of type is perhaps the most important single factor in legibility. Roethlein (38) stated that the legibility of a face undoubtedly increases directly with the increase in the size of a face. Freeman (16) found that 7 point type was read more rapidly than 11 point type and that the smaller type face seemed to allow somewhat greater scope of fixation, but that the number of fixations per second did not increase. However, Judd (26) found that the rate of reading and span of recognition are only slightly altered by doubling or halving the body size of a face. Parsons (34) is not opposed to condensed letters and thinks there may be an advantage in them. Huey (23) found that there is probably little to be gained by increasing the distance between letters (spacing of a word) beyond that which is used in the better printed books at the present time, although when the same number of words was printed in fewer lines of the same length and the same size of type, the material was read faster in the same proportion that the lines were decreased. Tinker and Paterson (44) obtained a speed of reading record for text which was set in roman lower case, italics and all capitals, and found the relative legibility of the type setting to be in the order named.

Leading: Noticeable differences have been found in extremely wide typographical settings in size and leading. Bentley (4) investigated Monotype News Gothic set in 12 point type and photographic reductions to correspond to 9 point and 6 point. He used as a criterion the time taken to read a given amount and concluded that 7 point leading for 12 point type, with corresponding equivalents in the smaller faces, resulted in the greatest savings, averaged at the rate of 4 minutes per hour. Typical newspaper reading conditions, however, were not present during the testing. Huey (23) found that leading was of no value when the size of type was less than 1.5 mm. for lower case nonprojectors. Parsons (34) concluded that the advantages of leading are almost negligible. Kerr (28) believed that legibility appeared to increase with leading but probably decreased as soon as one line above or below came within the small fixation area of distinct vision. Two

opposing results are to be found on the question of leading in the studies of the legibility of the telephone directory. Baird (2) concluded that within certain limits, legibility varied directly with the thickness of leading. Lyon (30) subsequently found, yet without reference to Baird's work, that there was no measurable difference between 6 point solid and 6 point type set on a 7 point base (leading 1 point), but that 1 point leading had greater aesthetic appeal and was generally preferred.

Length of Line: The width of a newspaper column today is standardized. However, earlier experiments confirm certain tendencies toward this standardization. Dearborn (13) found that 75 to 85 mm. or 3¾ inches, which is about ⅓ longer than daily newspapers, combined a good many advantages with a gauge of lower case n of 1.5 mm. The shorter lines are probably better than the longer lines because words in the line below are more likely to have a sense connection while otherwise they distract attention. Ruediger (39) favored a maximum of 75 to 85 mm., no matter what the subject matter. Huey (23) preferred short rather than long lines with approximately 66 mm. as the optimum and concluded that the line-length of modern newspapers are perhaps as near the optimum (for books too) as can be determined.

Margin: Like the length of the line, margins are standardized in newspapers. Legros (29) stated that a white space should logically form the boundary between adjacent areas of print. He favored a slightly wider margin where a metal rule would otherwise be considered necessary. However, the elimination of the dividing rule is not practical in newspaper publishing, as discovered by the New York Morning Telegraph, which attempted this recently.

Paper: The Report of the Committee of Illuminating Engineering Society (37) states that paper should possess a high (diffuse) reflecting power and little gloss. Any glazing is undesirable. Glossy inks have lower diffuse reflecting powers than the mat inks, hence the latter are blacker and present a greater contrast with the paper. Lyon (30) found no measurable difference between white papers and four shades of buff colored newsprint, although white paper had the greatest aesthetic appeal.

Many typographers have judged the fitness of the printed page by its aesthetic appearance and its historical origin. Morison (31), among others summarizes as follows: (*a*) the essential form of the letters should correspond with those that have been handed down to the reader historically; (*b*) the composition of the letters should be agreeable to the words; (*c*) the lower case letters should be homogeneous in line and color; (*d*) the upper case gauge should be at most slightly less than that of the lower case ascenders. The capitals must not be conspicuous and "spot" the page; (*e*) there should be no extreme contrast of the thick and thin strokes of the letters.

The approach to the various problems have been varied and the technique is open to question. As inferred, Pyke (35), after using a tachistoscopic method, doubted whether the differences in legibility found in the laboratory experiments hold in ordinary reading. Gilliland (17) recorded the reading rate and eye movements of subjects who read two kinds of material, ranging in typesetting from 6 to 54 points and 3 to 90 points paired with a 12 point control. Results were not significant between the limits of 6 to 36 points, although the rate of reading became somewhat slower with type larger than 18 point.

Empirical judgments of some writers are based only on a criterion of aesthetic value. Historical surveys, such as Blackhurst's findings (5) and Updike's review (45), present a picture of the attempt of the printing trade to solve the problem. Manufacturers of type reflect the feeling that the most legible type is the one of which the greatest quantity is sold over the longest period of time.

Experiments such as those of Dearborn (13) and Huey (21–23) have been conducted primarily with the object of analyzing the reading process. The work of Baird (2) and Lyon (30) on the telephone directory represents tests of speed and errors in finding a name and exchange number. Gray (18) studied the performance in oral and silent reading through the analysis of eye movements on a group of subjects of various intellectual atainment. Buswell (8), with a somewhat similar method and group of subjects, investigated the span of recognition for printed matter, the rate of recognition for perception units and the regularity of perception along the line of print. Judd and Buswell (27) had subjects read passages of various degrees of difficulty and familiarity, and found significant variations of reading performance in adjustment to the change in the content of the passages.

The dangers of speed reading has been discussed by Buswell (8), Judd (26), O'Brien (32), Pyke (35), and others. Similarly, the merits of loud and silent reading have been weighted. Inasmuch as silent reading forms practically the whole of adult reading in normal life, it is preferred to oral reading for test purposes, although it presents test difficulties. Maximum and normal reading has been compared by Schmidt (40), Whipple and Curtis (46) and others. Reading at the normal rate approaches the typical reading conditions, and if controllable for test purposes has a corresponding advantage.

In conclusion to this historical review of legibility, it may be stated that many conflicting and diametrically opposing results have arisen out of experimental evidence from a variety of sources. This may be explained by the nature of the observations and the limited experimental data upon which conclusions have been drawn. To this fact may be added the great number of empirical statements that have been made by printers and others who have been interested in typography and who

have judged factors of legibility by indefinite criteria. The approach to the problem should be wholly scientific and the emphasis of the psychological aspect cannot be overlooked.

DEFINITION AND MEASUREMENT OF LEGIBILITY

The term "legibility,"[1] as used in this paper, refers to the ease and accuracy with which the type setting can be read under ordinary conditions. Legibility is a factor in reading, and reading is the process of getting meaning from the printed page. Therefore, legibility may be defined as reading with ease and accuracy, where it is understood that meaningful reading takes place. This usage differentiates the term from perceptibility or recognizability, with which the experiment was not concerned.

Legibility has to do with reading in the normal and accepted manner. Thus the reader must sit in an upright position with good illumination falling upon the reading material which is held at a normal distance for accommodation to the eyes. Conditions should be typical of the actual reading situation and should be free from mechanical impositions which might tend to distract the reader.

The measurement of legibility[2] under these conditions is the

[1] Pyke (35) observed that four times as many writers have measured legibility as have defined it. The criterion of legibility was used by 36 writers in 98 years (1827–1924), whereas only 9 writers in 44 years (1878–1921) have defined legibility. Three out of every four writers have attempted to measure something the exact nature of which they have not paused to examine.

[2] Pyke (35) states that thirty-six writers have experimented and argued with fifteen different criteria. These are: "Measurement by speed of reading (by the time threshold and the amount read), the distance threshold (direct and peripheral), 'eye span,' 'illumination threshold' focus threshold, fatigue, number of eye focuses, number of eye refixations, regularity of eye-movements, reading rhythm, 'legibility coefficient,' 'specific legibility,' size of letters, by 'judgment of the trained human eye,' and by aesthetic merit. These criteria have been used indiscriminately under a variety of conditions; namely, in silent and in loud reading, at normal and at maximum speeds, using words (isolated and grouped) and letters (isolated and grouped), with daylight, oil, and gas-light, electric light and flash-light, with continuous reading,

rate of silent,[3] mature reading. Meaningful reading must take place and be checked for accuracy. The continuous text may be measured by the amount which has been read. Inasmuch as the reading material may be set in large or small type and the leading between the lines may increase the space that the eye covers in gaining meaning, the reading area is not a criterion of legibility. The measurement can be made logically by counting the number of words which can be read in 60 seconds. This measurement assumes that the writer of the reading material has been concise and economical in the choice of words, but even if the writer has been verbose the eye must transverse the wordy phrases to gain meaning. The words, by themselves and in combination, remain the basis of the perceptual units from which meaning is derived. The normal silent reading rate should give a close approximation of the legibility of the type setting.

PRELIMINARY CONSIDERATIONS

As a preliminary approach to the problem, 106 readers were given a comparative test of reading material set in two different newspaper body faces. Reader judgments were secured, but as the test was not given under standard conditions, its usefulness served only as an introduction to the problem at hand.

For practical assistance and standardized typesetting, I am indebted to Mr. Robert McLean, Vice-President of the Philadelphia Evening Bulletin, who not only appreciated the theoretic problem but offered coöperation greater than could be expected from any other source. At the time of my first contact with Mr. McLean, a change in the body face of the news columns of the page was contemplated.

---

and brief exposures varying from one-thousandth of a second to one-tenth, and contrived with tachistoscopes of four different sorts."

[3] An experiment by Schmidt (40) reveals the physiological basis of the superiority of silent over oral reading in point of both speed and comprehension. O'Brien (32) observed: "Silent reading is as a rule a more economical and effective instrument for the assimilation of the thought of a passage than oral reading."

The immediate cause of the change was the lack of legibility in the type used. This was determined by individual complaints on the part of readers, particularly noted when the dry mat for stereotyping was introduced. The dry mat or matrix may be contrasted with the wet mat which requires heat at a relatively lower pressure to form the matrix. The dry mat on the contrary requires much greater pressure (6 tons to the square inch) to form the impression on the papier-mache which is done by rolling the mat across the type form. The introduction of new type faces might not have been necessary except for this mechanical process. In the older type face, the more delicate hair lines and serifs broke down under several dry mat impressions. Multiple impressions were necessary to supply five casting rooms and a battery of 32 presses. To illustrate the manner in which a character will break down, let us take the letter O. The top and bottom of this letter, in most designs, is a hair line. The break down of this gives the appearance of two parenthesis marks. Again, a period may be battered so that it does not print an impression. The sum total of these faulty characteristics produced illegible reading and caused complaints by the readers.

The specifications for a new type face required heavier hair lines. The practical problems of stereotyping and of high speed printing demanded this. Furthermore, a similar type design was preferred because of the reading habits of the subscribers to the paper. After considerable study, it was thought best to have a type face similar to the Roman face most often used in school books. Childhood habits formed by reading this style of type face would persist, with the adult newspaper reader. No type face which was a radical departure from this specification was considered.

One rigid specification was that no type face would be accepted which was bastard within the range of type sizes customarily set for the body of a news article. This is illustrated in the cutting of type faces. In some type faces, it is necessary to redraw the designs for extreme variations in size if certain characteristics of a line of type is to be kept. This is

usually accomplished by extending or condensing letters and by changes in the ascenders and descenders of the characters.

The selection was finally narrowed down to five type styles which were set up for printing. Dry mats were made, which in turn were cast and printed according to the conditions under which they would be used. Of these five styles, the old type of the Bulletin was one. Further study proved that, of the remainder, only two styles were outstanding, as the characteristics of the others fell within these two styles. The conservation of newspaper printing space is important, and these two remaining styles offered variations in fractions of a point; namely, $6\frac{1}{2}$ and $6\frac{3}{4}$ point.

The final selection of the Intertype Ideal News Face (24) was decided upon by the publisher after all sizes were enlarged by photo-engraving and minutely compared with one another. The rejected type style proved to be bastard in the sizes larger than $6\frac{1}{2}$ point as the whole alphabet was condensed or distorted in order that the total alphabet should not be longer than in the $6\frac{1}{2}$ point size. However, the rejection was based on the additional fact that the larger type sizes seemed to "dance" before the eye if leading were not used.

Reading material for the experiment, recorded in this paper, was carefully considered. The selection of different material for the setting of each type size was necessary in order to avoid the influence of familiarity. Outstanding news stories of the year which had been published in various papers throughout the country were considered. After these had been compiled it was found that the style of writing and interest in the material varied to such an extent that it was not usable. Aesop's Fables, essays by Emerson, Schopenhauer and others were considered and rejected because of lack of similarity to newspaper stories.

Three factors were to be considered: (*a*) the material must be by the same author in order to avoid variations in the style of writing; (*b*) reading interest in the material must be sufficiently great to stimulate the normal reading rate; (*c*) the selections must be on the same level of reading difficulty.

The experience of editors of the Evening Bulletin was helpful in this regard. Over a period of twenty-five years it had been observed that dialect stories and stories in foreign background were less popular than stories of American life. A suitable story of common interest was found in Charles A. Lindbergh's "We" which had run in serial form. From this story sixteen incidents of approximately the same level of continuous interest and reading difficulty were selected.

TEST MATERIAL AND TYPE SETTING

The selection of reading material which was believed to contain items of approximately equal interest and difficulty follows:[4]

Column 1. Lindbergh's birth and family history.
Column 2. His educational background.
Column 3. A barnstorming trip with E. G. Bahl.
Column 4. Barnstorming with Lynch in Kansas.
Column 5. Flying his own ship.
Column 6. First cross country flight.
Column 7. Landing at Meridan in "Jenny."
Column 8. Storm conditions in flying.
Column 9. A negro as a paid passenger.
Column 10. Flight from Maben to Lake Village.
Column 11. Bud Gurney's parachute incident.
Column 12. Skakopee in a cloudburst.
Column 13. Information about Army Air Service.
Column 14. Flying in Nebraska.
Column 15. Flying cross country again.
Column 16. Description of aeroplane parachute.

The above headings give a brief description of the nature of the reading material. The style of writing permitted a blending of incidents with one another in a manner that allowed reading

[4] Among those who agreed with my personal judgment of the equality of the reading material were Dr. S. W. Fernberger and Mr. R. McLean and others who read the material for this purpose. Of unsolicited opinions, many of the individuals whose test results are included in the experiment stated that they believed the material to be uniformly equal in reading interest and difficulty. This fact was subsequently substantiated by inspection of the original copy as it had been prepared for printing in the different type sizes and leadings.

with continuous interest at the beginning of every column without having completed the reading of the previous column. Paragraph indentations followed the reading context as written by the author, except in one or two instances where an especially long paragraph could be divided logically in order to equalize column paragraphs. The phrase "Check mark" described under procedure, appeared at various distances from the top of the column but within a distance that the slowest reader would find the words before time was called. Only one such instruction appeared to a column.

The material was set in full newspaper page size according to the mechanical specifications of the Evening Bulletin:

Width of type page	17½ inches
Depth of type page	21 3/4 inches
Single column width	2 1/12 inches
Columns to page	8

At the top of the page in approximately 1½-inch depth across 8 columns were printed:

TEST FOR LEGIBILITY OF TYPE

Wait for Instructions Before Reading Any Column

At the bottom of the page in approximately 2 inch depth across 8 columns was printed:

WAIT FOR INSTRUCTIONS BEFORE ANSWERING THESE QUESTIONS

And in three divisions across the page were set:

Name.....

Check Sex: Male Female

Age.....

Do you believe that you have normal vision or corrected vision with glasses?

Yes No

If not, what is the nature of your eye defect?.....

Which of the above columns seemed easiest for you to read?

Column No. .....

Why did it appear the easiest to read?.....

.....

Each of the eight columns across the page were divided by column rules in newspaper fashion. Each newspaper column was divided into two 8-inch columns of reading matter headed

Column 1 Column 2

etc., in approximately a $\frac{3}{4}$-inch space, below which appeared reading matter, separated by a hair line rule. In this manner all reading matter was separated. Columns 1 to 8, inclusive, appeared on the upper half of the page and columns 9 to 16, inclusive, on the lower half.

The length of each line was $12\frac{1}{2}$ ems ($2\frac{1}{12}$ inches) butted next to a double bevel hair line rule with a 6 point base, leaving a margin of approximately 3 points, with the exception of the wider margin on the outside of columns 1, 8, 9 and 16. The characters adjacent to either margin were flush against the column rule with intervening letter spacings accompanied by the band space of the linotype machine.

The reading material was set in the following type sizes:

COLUMN	GROUP TEST A	GROUP TEST B AND C
1	$6\frac{3}{4}$ on 7	8 on 9
2	$6\frac{1}{2}$ on 7	6 on 7
3	$7\frac{1}{2}$ on 7	7 on 8
4	6 on 7	$6\frac{1}{2}$ on 7
5	$6\frac{3}{4}$ on 7	6 on 6
6	$6\frac{3}{4}$ on $6\frac{3}{4}$	$6\frac{3}{4}$ on $7\frac{1}{2}$
7	$6\frac{1}{2}$ on $6\frac{1}{2}$	8 on 8
8	6 on 6	$6\frac{3}{4}$ on 7
9	8 on 8	$6\frac{1}{2}$ on $6\frac{1}{2}$
10	$6\frac{3}{4}$ on $7\frac{1}{2}$	$6\frac{3}{4}$ on $6\frac{3}{4}$
11	$6\frac{1}{2}$ on $7\frac{1}{2}$	$6\frac{3}{4}$ on $7\frac{1}{4}$
12	7 on $7\frac{1}{2}$	8 on $8\frac{1}{2}$
13	7 on 8	7 on 7
14	8 on 9	$6\frac{3}{4}$ on 7
15	8 on $8\frac{1}{2}$	7 on $7\frac{1}{2}$
16	$6\frac{3}{4}$ on $7\frac{1}{4}$	$6\frac{1}{2}$ on $7\frac{1}{2}$

Explanation: $6\frac{3}{4}$ on 7, refers to $6\frac{3}{4}$ point Ideal type on a 7 point base, or leaded $\frac{1}{4}$ point; 7 on 7 refers to the size of type on the same base with no leading.

The point measurement used in this experiment is a modification of the French System which was adopted by the United States Type Founders' Association in 1886. As 996 points was the nearest metric measure of 35 centimeters, the point measurement in inches is mathematically .0138648 inches, usually spoken of as approximately $\frac{1}{72}$ of an inch. The equivalents of the type sizes used in the experiment and defined by the American Type Founders' Association are:

6 point	0.083 inch
7 point	0.09683 inch
8 point	0.1106 inch
9 point	0.1245 inch

However, the measurement used on the Mergenthaler Linotype machines which were used in setting this material in the Bulletin composing room, are as follows:

6 point	0.084 inch
6½ point	0.091 inch
6¾ point	0.0945 inch
7 point	0.098 inch
8 point	0.112 inch
9 point	0.126 inch

The designation of Group Tests A, B and C is used for convenience in reference to both test subjects and material. Each group test contains an equal number of subjects participating in the experiment. Group Test A represents the original reading material set in type sequence of random choice. In Group Test B the type setting of reading material has been reversed. This was accomplished by reversing extremes of the reading rate after the results of Group Test A had been secured. Rank order 1 of the average reading rate was set in the type setting of rank order 16, and so on until rank order 8 was changed with rank order 9. Group Test C is identical with Group Test B except for a difference in subjects and point of time in which the test was conducted. The reading material in Group Tests A, B and C is the same throughout the tests inasmuch as the order of reading for each column and the context for each column has been held constant.

From the original page form of the type setting a dry mat was made, and in turn a stereotype was cast which printed on a high speed press. The ink was Charles E. Johnson's black High Speed News Ink, and the paper grade was blue white newsprint of a 32 pound base. The same material was printed on both sides of the paper in order to produce the similar grayness of a newspaper page.

SUBJECTS FOR EXPERIMENT

Three groups of 100 individuals each took part in this experiment. The performance of these groups are designated under the headings, Group Tests A, B and C. The subjects were male[5] college students in the Sophomore and Junior years, all students or prospective students in an Advertising course at the Wharton School of Finance and Commerce, University of Pennsylvania. The students whose tests were recorded were naïve relative to typographical information. Those who were tested in the course in advertising had not at the time of the test studied typography.

By using groups of 100 individuals, it was expected that the variation of individual differences and errors (both physical and temperamental) would be cancelled by the law of average and that the result would present an accurate picture of the ordinary reader. Physiological habits in reading, such as the perceptual span, the fixation pause, eye movements, refixations, the return sweep and inner speech, were considered as constant factors that would be transmitted with equal weight to each of the columns that were read by the individual subject. Results from individuals who knew that they had defective eyesight and who lacked corrected vision by use of glasses, were discarded.

Every student had been admitted to the University by means of a psychological test or the Scholastic Aptitude Test so that normal intelligence may be presumed. While the intelligence of the group may be considered higher than the average newspaper reader, the known intelligence of this group permitted a

[5] Brooks (6) states that in silent reading sex differences are not marked.

better check on the accuracy of mature reading than would be possible if a wider range of intelligence were present. Only the results of native-born English reading students were accepted.

CONDITIONS OF TESTING

The experiment was conducted over a period of two college years. Group Test A was made in the spring months of 1928 at the close of one college year, and Group Tests B and C were made in the fall months of 1928 at the beginning of a college year. The experiment was held during a classroom hour of 50 minutes. Sixty seconds was the allotted reading time to each column with a greater rest period between readings to relieve fatigue.

The reading took place between 8 o'clock in the morning and 4 o'clock in the afternoon, with the bulk of the reading appearing nearer noonday. Natural daylight was used which was reinforced by electric ceiling lights. Walls and ceilings of the rooms were painted cream. It is believed that the illumination approximated the reading conditions in which the average newspaper is read.

Typical reading conditions were present. It was expected that each subject would read naturally and in his normal manner. Instructions were repeated three times in order that there be no misunderstanding. Personal comfort was extended to the subject in permitting any reading position that was natural. A large arm rest upon each chair permitted ease of writing. The paper was held in the manner in which the reader customarily accommodated his vision to the page.

Silent reading was followed throughout. The procedure emphasized the fact that it was not a speed test and that meaningful reading must take place. Caution was exercised throughout the experiment toward this end. The group instructions were given under the heading of Procedure incorporates the test method.

PROCEDURE

The test reading was undertaken in sections of ten to fifteen individuals. The results of each individual were combined

into Group Tests A, B and C as previously indicated. The size of the section was small enough to facilitate careful direction of each individual so that the instructions did not miscarry. None of the individuals expected a test during the hour in which it was given and great care was exerted in order to avoid the appearance of a competitive test. The following instructions were given:

"I am going to ask you to do a little reading today. I have here excerpts from Lindbergh's story 'We.' You will find that the reading is interesting, but I do not want you to consider it as a test since I am not concerned with your individual ability but with the average ability of a large group of individuals. Please follow my instructions implicitly. At the outset I am going to ask you to take alternative seats. You will find this more comfortable as you will not be crowded by your neighbor. Furthermore, I must ask you not to look at your neighbor's paper at any time during the period in which we shall read the articles which I have had specially printed. You might consciously or unconsciously enter into a competition with him. However, at the end of the period I shall give you an opportunity to make comparisons if you wish."

(At this point the test papers were held up to show the group.)

"On this page are sixteen columns, each containing a different excerpt of Lindbergh's story 'We.' Wait for instructions before reading any column. I am going to ask you to read each column normally as you would read a newspaper, but be sure you read for content.

"The reading instructions for each column will be given like this: 'Get ready to read column 1; Ready; Read!' These instructions correspond to the instructions given a track athlete, but remember you are not running a race and I want you to read *normally* and for *content*. The instructions are so given that you will have the proper attitude and a similar attitude before reading each column.

"At the word 'Read' you will start to read normally and for content. In order that I may be sure that you have read for content, I have inserted the words "(Check mark)" in parenthesis within the columns. (At this point the words (Check mark) are placed on the blackboard. These words have nothing to do with the reading matter and may appear anywhere within a sentence, paragraph or column. Whenever you see these words, put a "check mark" (√) through the words in this manner. (Here a check mark (√) was placed through the words on the blackboard.)

"After following this instruction, continue to read until I say: 'Stop.' Put a cross through the last word read. (Here a cross (×) was placed on the blackboard.) Be sure you put a cross through the *last word read*,

whether this word is at the first of a sentence, the middle of a sentence, or at the end of a sentence."

"When this is done, be sure to look up and away from the paper. By no means continue to read. If you are interested in the story, I shall give you an opportunity to come back to it later. I want you to look away from the paper toward the front of the room or out of the window. I ask you to do this as relaxation. If you raise your arm once, you suffer no fatigue, but if you raise your arm ten or more times, you may notice it. Similarly you may not notice eye fatigue in reading a short time, but I want to make sure that you read each column under the same circumstances, so please look away as soon as you place a cross through the last word read. I shall then ask members of the class what they have read."

(At this point the papers were distributed to the group.)

"Please do not look at the paper that I am giving you. It is printed the same on either side of the sheet, but do not look at it or fill out anything until I give you further instructions. Remember, you have nothing to gain by trying to look at the paper ahead of instructions, so please curb your curiosity. If you do look, it is like cheating in solitaire —there is no point to it."

After the papers were distributed, the substance of the foregoing instructions was repeated and questions were asked in order to clear up any vague point. Again the instructions were repeated for the third time:

"Let me repeat briefly. You are going to read column 1 in this manner. I shall say: 'Get ready to read Column 1; Ready? Read!'—at which time you will start reading, being sure that you read normally and for content. When you see the word (Check Mark), follow the instructions (this is emphasized by pointing to the blackboard where the checkmark (√) appears through the words.) This done, continue to read until I say 'Stop.' Put a cross through the last word read. Are there any questions?

"There is one thing I might add. The time given you to read is so regulated that I do not expect the fastest reader of this group to finish the column, but should I be mistaken, will you note the fact that you have read the column in less than the allotted time. Should this be the case, of course do not read into the next column.

"Are the instructions clear?"

"Get ready to read Column 1; Ready? Read!"

The stop watch was pressed simultaneously with the word "Read." During the sixty second reading interval, I watched carefully the individuals who were reading. Distractions were noted and a caution was given at the end of the reading. In

the case of an individual who had a severe cold and who was forced to take ten or fifteen seconds out of one reading period in order to clear his head, his results were discarded.

At the end of the sixty second reading period, I was insistent that the instructions were followed faithfully. Reading stopped and the individuals all looked away from their papers.

I then asked: "Mr. Blank, where was Lindbergh's father born? Mr. So and So, give the incident described in this column." If evidence was found that the reading was faulty, the result was discarded. I then looked at each paper to see that the check mark was properly made and asked for questions on the procedure.

Before reading Column 2, the group was again cautioned that they should read normally as they would read at home without these conditions being imposed upon them.

The rest period between columns was not timed, but as the test took the full class hour, the rest period always exceeded the reading period. The remaining columns were read under the standard conditions. At the end of each reading the questions varied according to content and were also varied in the manner asked. Thus a question might be asked: "The name of Lindbergh's friend was: Smith, Lynch, Custer. Which was it?" These names had appeared in the reading, but Custer was a battlefield and Lynch was Lindbergh's friend. Again, "What type of aeroplane was used?" "Through which states did Lindbergh pass?" or "In what town did Lindbergh stop?" These specific questions were reinforced by a further statement of the incident described in each column.

At the completion of the reading of the sixteenth column, the individuals filled in the personal data relative to their age and the nature of their vision as they knew it, and made a judgment of which column seemed easiest to read and why. No clue to this latter information was given to them.

At the end, the papers were collected individually and each paper was carefully inspected to see that the instructions had been properly filled out. When known eye defects were discovered, the results of these individuals were discarded.

Furthermore, results of a Chinese and a French Canadian student were discarded because English was not their native tongue, although from casual observation their results appeared to be on a par with the others.

In obtaining the results, the following additional tests had been made, but were not used: 18 were discarded because of known eye defects; 18 were tested in the second and third experiments whose results had appeared in the first, and these results were not included because of the influence of familiarity; 23 were discarded due to errors in following instructions or reading the allotted column in less than 60 seconds; two were foreign-born individuals. Little uniformity was found among those who read the allotted material in less than 60 seconds, and the faster reading of these individuals probably resulted from a competitive spirit, skimming, lack of interest and greater interest in material, anticipating the starting signal, and similar factors which may be classed, as a whole, as errors in following instructions. No one read over half of the sixteen columns in less than the allotted time. Of this discarded group few overread more than one of the sixteen columns. The discarded papers were not tabulated.

RESULTS AND INTERPRETATION OF DATA

The 300 individual results were compiled in three groups, Group Tests A, B and C. A total of 1,318,343 words were read throughout the 4800 separate columns in the combined reading time, exclusive of rest periods, of 80 hours. While it is possible to include in the tables only the arithmetic means of the groups, anyone who may be interested may secure photostatic copies of the original data. A study of the individual records shows variations from the average but these variations are for the most part within the limitations of sampling, and the tendency of the reading rate may be well expressed by the arithmetic mean. Part I includes the results of the direct measurement of the number of words which were read; Part II describes the readers' preferences and opinions of the legibility of the type setting.

PART I

Table 1 gives a comparison of the type setting with the average number of words read in 60 seconds. The table includes Group Test A, B and C, each containing the average results of 100 readers. Besides the reading rate, the table includes the rank of the arithmetic mean, the reading order of the context,

TABLE 1

Comparison of type setting with reading rate

SIZE OF TYPE	LEADING OF TYPE	GROUP TEST A				GROUP TEST B				GROUP TEST C			
		Av. number words read in 60 sec.	Rank order of arith. mean	Reading order of context	Readers' preferences	Av. number words read in 60 sec.	Rank order of arith. mean	Reading order of context	Readers' preferences	Av. number words read in 60 sec.	Rank order of arith. mean	Reading order of context	Readers' preferences
6 on 6		289.22	5	8	—	258.98	12	5	—	264.89	12	5	—
	6 on 7	274.18	11	4	1	267.50	10	2	1	273.22	8	2	—
6½ on 6½		261.96	16	7	—	289.51	2	9	4	290.02	2	9	1
	6½ on 7	285.96	6	2	—	256.63	13	4	—	263.16	14	4	1
	6½ on 7½	302.47	2	11	2	251.83	14	16	—	259.38	15	16	3
6¾ on 6¾		278.47	8	6	—	275.43	7	10	—	272.67	9	10	1
	6¾ on 7	265.14	14	1	3	281.90	5	8	2	287.80	3	8	1
	6¾ on 7	273.24	12	5	—	283.31	3	14	—	277.11	7	14	1
	6¾ on 7½	277.15	9	10	1	269.17	8	6	2	281.19	4	6	8
	6¾ on 7¾	264.28	15	16	5	302.32	1	11	10	305.29	1	11	10
7 on 7		272.85	13	3	2	282.28	4	13	5	280.61	5	13	5
	7 on 7½	282.84	7	12	—	268.07	9	15	—	270.29	10	15	—
	7 on 8	290.85	4	13	11	251.36	15	3	7	263.76	13	3	11
8 on 8		306.19	1	9	30	262.16	11	7	7	266.18	11	7	4
	8 on 8½	277.02	10	15	13	277.67	6	12	32	277.12	6	12	24
	8 on 9	292.94	3	14	32	239.21	16	1	30	238.65	16	1	30

and also (this concerns Part II of the experiment) the readers' preferences as to which type face was judged easiest to read.

Table 2 contains the same data shown in table 1 with the exception that the reading order of the context has been kept in sequence and hence is a comparison of context with the reading rate.

The information contained in tables 3, 4 and 5 may be brought forward as evidence of the reliability of the arithmetic mean. Table 3 gives the means of the reading rates for each of the sixteen reading samples in Group Test A. It gives also the standard deviation of each sample and the upper and lower limits of the range in standard deviation on either side of the

TABLE 2

Comparison of context with reading rate

READING COLUMN NUMBER	GROUP TEST A				GROUP TEST B				GROUP TEST C			
	Av. number words read in 60 sec.	Rank order of arith. mean	Reading order of context	Readers' preference	Av. number words read in 60 sec.	Rank order of arith. mean	Reading order of context	Readers' preference	Av. number words read in 60 sec.	Rank order of arith. mean	Reading order of context	Readers' preferences
1	265.14	14	1	3	239.21	16	1	30	238.65	16	1	30
2	285.96	6	2	—	267.50	10	2	1	273.22	8	2	—
3	272.85	13	3	2	251.36	15	3	7	263.76	13	3	11
4	274.18	11	4	1	256.63	13	4	—	263.16	14	4	1
5	273.24	12	5	—	258.98	12	5	—	264.89	12	5	—
6	278.47	8	6	—	269.17	8	6	2	281.19	4	6	8
7	261.96	16	7	—	262.16	11	7	7	266.18	11	7	4
8	289.22	5	8	—	281.90	5	8	2	287.80	3	8	1
9	306.19	1	9	30	289.51	2	9	4	290.02	2	9	1
10	277.15	9	10	1	275.43	7	10	—	272.67	9	10	1
11	302.47	2	11	2	302.32	1	11	10	305.29	1	11	10
12	282.84	7	12	—	277.67	6	12	32	277.12	6	12	24
13	290.85	4	13	11	282.28	4	13	5	280.61	5	13	5
14	292.94	3	14	32	283.31	3	14	—	277.11	7	14	1
15	277.02	10	15	13	268.07	9	15	—	270.29	10	15	—
16	264.28	15	16	5	251.83	14	16	—	259.38	15	16	3

mean. In normal distribution 68 per cent of the cases will fall within a range of one standard deviation above or below the arithmetic mean. An examination of the seventh column of the table shows that in the sixteen samples from 62 to 70 per cent of the total cases fall within this range. The variations in the percentage of cases within the range of one standard deviation above or below the mean may be explained in terms

of the probable error of the several standard deviations. Furthermore a measure of skewness by the position of averages method showed a very small amount of assymetry. The maximal amount for any sample was –8 per cent. Hence, while this is not a complete test for normality it may be assumed

TABLE 3

Reliability of data

Group Test A

READING COLUMN	ARITH. MEAN*	S.D.	A.A. – 1 S.D.*	A.A. + 1 S.D.*	COEF. OF VAR.	PER CENT TOTAL CASES INCLUDED WITHIN RANGE OF 1 S.D. ON EITHER SIDE OF MEAN
1	265	55.44	210	320	20.9	69
2	286	51.60	234	338	18.0	61
3	273	45.26	228	318	16.6	62
4	274	50.85	223	325	18.5	71
5	273	43.28	230	316	15.9	64
6	278	50.99	227	329	18.3	67
7	262	44.03	218	306	16.8	67
8	289	57.87	231	347	20.0	70
9	306	53.94	252	360	17.6	62
10	277	56.47	221	333	20.4	67
11	302	54.71	247	357	18.1	69
12	283	59.13	224	342	20.9	68
13	291	49.19	242	340	16.9	70
14	293	51.31	242	344	17.5	68
15	277	50.73	226	328	18.3	65
16	264	49.17	215	313	18.6	69

* Decimal point omitted and fraction of 0.5 or more included as whole number.

that the distribution of the sixteen samples were approximately normal and that the variations from the average were those that could be expected in simple sampling. For the benefit of those who may be interested the coefficient of variation is given in the sixth column of the table. Tables 4 and 5 give the same information for Group Tests B and C.

Table 6 gives the correlations between the various tests by the Pearsonian coefficient formula. The correlations are all high. Group Test A represents the first test. Because of the possible influence of context, despite the fact that the material was carefully chosen, the type setting was reversed

TABLE 4

Reliability of data

Group Test B

READING COLUMN	ARITH. MEAN*	S.D.	A.A. − 1 S.D.*	A.A. + 1 S.D.*	COEF. OF VAR.	PER CENT TOTAL CASES INCLUDED WITHIN RANGE OF 1S.D. ON EITHER SIDE OF MEAN
1	239	50.25	189	289	21.0	67
2	268	46.78	221	315	17.5	65
3	251	43.12	208	294	17.1	68
4	257	45.02	212	302	17.5	71
5	259	47.77	211	307	18.4	68
6	269	54.03	215	323	20.1	67
7	262	43.46	219	305	16.6	62
8	282	57.18	225	339	20.3	66
9	290	52.58	237	343	18.1	66
10	275	64.14	211	339	23.3	66
11	302	61.36	241	363	20.3	70
12	278	61.04	217	339	21.9	64
13	282	53.62	228	336	19.0	73
14	283	60.47	223	343	21.4	70
15	268	56.02	212	324	20.9	77
16	252	47.42	205	299	18.8	74

* Decimal point omitted and fraction of 0.5 or more included as whole number.

for Group Test B and C in proportion to the extremes of the reading rate. This is equivalent to inverting the context with respect to the original type setting. In Group Test A, Column 9, set in 8 point type on an 8 point base, ranked first, and Column 7, set in 6½ point type on a 6½ point base, ranked

sixteenth. In Group Test B, Column 9 was set 6½ on 6½, and Column 7 was set 8 on 8. This change was subsequently followed by the reverse setting of Columns 2 and 14, 3 and 13, 4 and 12 until each column was completely changed to a different type size or leading. Group Test B and C represent the same material read by different individuals.

TABLE 5

Reliability of data

Group Test C

READING COLUMN	ARITH. MEAN*	S.D.	A.A. − 1S.D.*	A.A. + 1S.D.*	COEF. OF VAR.	PER CENT TOTAL CASES INCLUDED WITHIN RANGE OF 1S.D. ON EITHER SIDE OF MEAN.
1	239	63.50	175	303	26.6	66
2	273	65.11	208	338	23.8	59
3	264	56.72	207	321	21.5	64
4	263	57.02	206	320	21.7	72
5	265	61.11	204	326	23.1	66
6	281	61.86	219	343	22.0	64
7	266	54.30	212	320	20.4	69
8	288	73.40	215	361	25.5	66
9	290	63.61	226	354	21.9	70
10	273	70.14	203	343	25.7	64
11	305	62.51	242	368	20.5	64
12	277	68.80	208	346	24.8	68
13	281	62.32	219	343	22.2	68
14	277	66.36	211	343	24.0	67
15	270	63.11	207	333	23.4	70
16	259	61.55	197	321	27.7	72

* Decimal point omitted and fraction of 0.5 or more included as whole number.

Group Tests B and C were performed under identical conditions except as to point of time when the tests were made. With the exception of the extremes in the reading rates Group Tests B and C vary within the limits of simple sampling.

According to the independence condition of sampling, there should be no correlation between the results of the two tests. The coefficient of correlation is, however, 95 per cent, ranked either according to the size of type or order of context. Hence we must conclude that there is some association between the two tests. There are two possible causes of relationship: (*a*) Variation of type setting, (*b*) Variations of context.

In Group Tests A and C, the context was held constant and the order of type setting was inverted. Under these conditions 74 per cent direct correlation was found between the tests.

TABLE 6

Group test correlations

Comparison of Type Setting with Reading Rate	
Data from table 1	
Correlations between Group Tests	**A and B = −84.2**
	A and C = −79.8
	B and C = +95.0
Comparison of Context with Reading Rate	
Data from table 2	
Correlations between Group Tests	**A and B = +90.5**
	A and C = +74.0
	B and C = +95.0

Correlations were computed from the mean by use of the Pearsonian coefficient formula.

Inverting the order of context and holding the order of type setting constant we find 79.8 per cent inverse correlation. Since inverting the order of context has changed a high positive correlation to an equally high negative correlation, it follows that the order of context is the dominant characteristic in determining the association between the two tests.

The conditions for Group Tests A and B, and A and C, are the same except as to point of time. We find the conclusions reached by comparing Group Tests A and C to be substantiated by a comparison of Group Tests A and B. With order of context held constant and order of type setting varied we have

90.5 per cent direct correlation. Inverting the order of context and holding the order of the type setting constant we find 84.2 per cent inverse correlation. Again it follows that the dominant characteristic is order of context.

TABLE 7

Ratio of the actual differences of the mean of each sample from the mean of all the samples to the standard error of the differences

RANK ORDER OF ARITH. MEANS	D/S.E. of Diff.		
	Group Test A	Group Test B	Group Test C
1	3.3	3.9	3.6
2	2.7	2.6	1.9
3	1.6	1.6	1.5
4	1.4	1.6	.9
5	1.0	1.5	.9
6	.7	1.0	.4
7	.3	.6	.4
8	.4	.1	.0
9	.5	.3	.0
10	.5	.3	.3
11	1.0	1.1	.8
12	1.2	1.5	.9
13	1.2	1.8	1.0
14	2.1	2.5	1.2
15	2.4	2.7	1.6
16	2.8	4.1	3.7

Computed from formula on the reliability of the difference in terms of the standard error of the difference.

It is usually customary to take D/S.E. of Diff. of 3 as indicative of complete reliability. See "Statistcs in Psychology and Education" by Henry E. Garrett.

The method by which the experiment was conducted precludes the possibility of measuring the general relationship between the rate of reading and the type setting. The close correspondence between the positive correlation on the one hand and the negative correlation on the other, when the order of the context is inverted, argues fairly conclusively that no general association exists between rate of reading and the type setting.

This fact makes it necessary to turn to an investigation of the variations in the rate of reading for the individual samples.

In Group Tests A, B and C an examination of the variation in the rate of reading of each sample from the average rate of the 16 samples (table 7) shows twelve cases in which the actual differences were sufficiently in excess of the standard error of the differences to be significant. A further investigation (table 8) indicates that these significant differences are associated with the same context. Thus the context of the reading columns 9 and 11 gives reading rates outside the limits of simple sampling in all three group tests. This may be explained

TABLE 8

Significant differences in extreme samples

RANK ORDER OF MEAN	GROUP TEST A			GROUP TEST B			GROUP TEST C		
	Size of type	Context	D/S.E. of diff.	Size of type	Context	D/S.E. of diff.	Size of type	Context	D/S.E. of diff.
1	8 on 8	9	3.3	6¾ on 7¾	11	3.9	6¾ on 7¾	11	3.6
2	6½ on 7½	11	2.7	6½ on 6½	9	2.6	6½ on 6½	9	1.9
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
15	6¾ on 7¾	16	2.4	7 on 8	3	2.7	6½ on 7½	16	1.6
16	6½ on 6½	7	2.8	8 on 9	1	4.1	8 on 9	1	3.7

by the fact that these columns contained slight elements of humor and suspense, respectively, which give them exceptional interest value. The reading material in columns 1, 3, 7 and 16 was descriptive in nature, both historical and factual, but the factors which have caused these differences are not sufficiently in evidence to be determined. Likewise in the other columns no conclusion can be drawn from the relative interest or reading difficulty of the context.

Of the variables, type setting and context, context is of relatively greater importance than type setting. But as only one column is in perfect agreement among the three tests and that only as to rank order, which may be accidental, one may

inquire further as to the differences. The actual difference of the number of words read between the most read and least read columns gives this contrast:

	GROUP TEST A	GROUP TEST B	GROUP TEST C
Rank 1	306.19	302.32	305.29
Rank 16	261.96	239.21	238.65
Word difference	44.23	63.11	66.64

In reality this comparison only shows the similarity between the tests which the correlations already noted. However, it may give some indication of differences within the group. The average age of Group Test A is 20.55 years, Group Test B, 20.21 years, and Group Test C, 19.47 years. Also, Group Test A was given in the spring of the year and the other tests were given in the fall. The total number of words read throughout all columns is: Group Test A—449,476; Group Test B—431,733; Group Test C — 437,134. This factor indicates something other than the relative maturity of the subjects, as it is not coincident with it. The difference will be considered later under the general discussion of the experiment.

Possibly an interesting result, though shedding no light on the matter of size of type and leading as a factor of legibility, is found in the two columns set in 6¾ on 7. This duplication was originally inserted as a checking factor, being used in Group Test A in Column 1 and 5 where Column 1 was thought to be necessary as a "warming up" column. However, when the context was believed to be an important influence, the setting for the subsequent tests were reversed and the "warming up" column eliminated. Taken from Table Nunber I is the following comparison:

SIZE	GROUP TEST A	GROUP TEST B	GROUP TEST C
6¾ on 7	265.14 (14) 1	281.90 (5) 8	287.80 (3) 8
6¾ on 7	273.24 (12).5	283.31 (3) 14	277.11 (7) 14
Word difference	8.10	2.41	10.69

No significance can be attached to the above information since it has been shown that the influence of context has permeated throughout the experiment. The word differences might be considered as some indication of the relative difference in context among columns 1, 5, 8 and 14. However, this information serves no purpose in determining the relative value of size and leading for the experiment as a whole.

How futile this becomes is shown in these subsequent tables:

Relationship of type size, based on rank order

SIZE OF TYPE	GROUP TEST A	GROUP TEST B	GROUP TEST C
6 on 6	5	12	12
6½ on 6½	16	2	2
6¾ on 6¾	8	7	9
7 on 7	13	5	5
8 on 8	1	11	11

It has generally been assumed by those who work with type that legibility is directly proportionate to the size and leading, within limits. In this experiment the reversal of the type setting in Group Tests B and C from Group Test A is equivalent to inverting the context with respect to the original type setting. The facts do not show a proportionate increase in legibility as the type size increases but the reading rate appears to be influenced entirely by context.

In regard to leading, this table shows a similar influence:

Relationship of leading 1 point, based on rank order

SIZE AND LEADING	GROUP TEST A	GROUP TEST B	GROUP TEST C
6 on 7	11 −	10 +	8 +
6½ on 7½	2 +	14 −	15 −
6¾ on 7¾	15 −	1 +	1 +
7 on 8	4 +	15 −	13 −
8 on 9	3 −	16 −	16 −

For convenience of interpretaton, the plus and minus signs indicate improvement and retrogression in the rank order

when 1 point leading was added. Only two out of the five cases show any improvement with leading, and if it were not for the context factor, it might be concluded that leading is of little or negative value in regard to legibility. It is evident that there is not only a lack of agreement in the influence of type sizes, but also there is a lack of agreement in the influence of leading.

In five columns, the same type size appears either on the same base or leaded from $\frac{1}{4}$ point to 1 point. The effect of leading appears in tabular form below:

Relationship of leading same type face, based on rank order

SIZE AND LEADING	GROUP TEST A	GROUP TEST B	GROUP TEST C
$6\frac{1}{4}$ on $6\frac{1}{4}$	8	7	9
$6\frac{1}{4}$ on 7	14 −	5 +	3 +
$6\frac{1}{4}$ on 7	12 −	3 +	7 +
$6\frac{1}{4}$ on $7\frac{1}{2}$	9 −	8 −	4 +
$6\frac{1}{4}$ on $7\frac{3}{4}$	15 −	1 +	1 +

The few cases of the same single type size with leading do not warrant a separate statistical correlation as the few items will result in a large probable error. With one exception the effect of leading has inverted the order of the results of Group Test A, when the context was changed in Group Test B and C. If a similar comparison is taken from the entire experiment, the effect of leading as a whole over 16 columns of all three tests, 13 examples are found where leading has aided legibility and 20 where it has retarded the reading rate. These figures are cited only generally as no conclusions can be drawn from them since we have previously demonstrated the dominant influence of context.

CONTINUATION

For convenience in publication, it has been necessary to divide the paper at this point. For the purpose of continuity, it may be noted that Part II deals with the reader's preferences and opinions of legibility in contrast with the criterion of the reading rate. The discussion of both Parts I and II along with the general conclusions and bibliography must be left to the next issue of the Journal.

NOTE ON THE COMPARATIVE VALIDITIES OF JUDGMENTS OF INTELLIGENCE BASED ON PHOTOGRAPHS AND ON INTERVIEWS

ETHEL MORIWAKI

Territorial Normal and Training School, Honolulu Hawaii

The subjects used in this study were seven men and eight women students of the Territorial Normal School, Honolulu. Their ages ranged from eighteen to twenty-one years. They were specially selected to represent the full range of intelligence found in this institution, being chosen at regular intervals of five score points on the Thorndike Intelligence Examination.

The judges consisted of four prominent educators of Hawaii, the Assistant Director of the Territorial Division of Research, the Principal of the Experimental School, the First Grade Supervisor of the Normal Training School and the Director of Teacher Training of this Institution.

The photographs were all taken by the same person with the same camera in the same room, and were front views from head to bust. They were arranged alphabetically for presentation to the judges. Each judge ranked all the photographs without consultation.

Two weeks after the photographs were examined and judged, the fifteen subjects were asked to present themselves individually in a classroom for a personal interview with the judges. Each one was asked questions and given an opportunity to express himself in speech for about two minutes. At the end of the interview, each judge recorded his impressions of the subject's intelligence without conferring with the others. At the close of the interviews each judge ranked the subjects.

The average school grade for the first two quarters was computed for each subject.

Correlations were computed by the Spearman rank-difference formula. These correlations are given in the accompanying table. The four judges are designated A, B, C and D when judgments of the photographs are under consideration, and A′, B′, C′ and D′ when considering interviews. The total judgments of the last two columns were obtained by summing the four ranks for each subject and ranking these sums. (See table 1.)

From the table we find that the average correlation between Thorn-

dike score and judgments based on photographs is .28; between Thorndike score and judgments based on interviews, .56; between average grade and judgments based on photographs, .45; between average grade and judgments based on interviews, .44; and between judgments based on photographs and judgments based on interviews, .49. The average intercorrelation between judgments based on photographs is .48; between judgments based on interviews, .59; and between one set of judgments based on photogrpahs and another based on interviews (exclusive of the four above noted in which the same judge made both judgments), .29.

TABLE 1

Coefficients of correlations

	GRADE	PHOTOGRAPHS				INTERVIEWS				TOTAL JUDGMENT OF PHOTOS	TOTAL JUDGMENT OF INTERVIEWS
		A	B	C	D	A′	B′	C′	D′		
Thorndike.....	.79	.40	.22	.36	.15	.51	.70	.45	.58	.41	.63
Grade.....		.56	.42	.41	.39	.32	.55	.42	.46	.45	.48
A.....			.39	.58	.24	.57	.32	.19	.63	.70	.43
B.....				.67	.48	.34	.40	.16	.46	.72	.42
C.....					.51	.40	.16	.28	.35	.76	.35
D.....						.13	.08	.15	.69	.66	.16
A′.....							.64	.45	.85	.51	.87
B′.....								.51	.77	.30	.77
C′.....									.34	.18	.62
D′.....										.52	.90
T.....											.38
W.....											

The significance of these results may be gauged by the fact that with a sample of 15 drawn from an uncorrelated population, ρ's as high as .28 would arise about once in six trials as a matter of chance, as high as .52 about once in 40 trials, and as high as .62 about once in 160 trials. The probability that 45 such correlations would all be positive, however, is practically zero.

Hence we may conclude that judgments of intelligence based on photographs are definitely though slightly better than chance; while judgments based on interviews are noticeably better, and in fact seem to have here some real validity in separating individuals over a wide range of intelligence.

A COMPARATIVE STUDY OF PSYCHOGALVANIC AND ASSOCIATION TIME MEASURES

A New Psychogalvanic Apparatus

STARKE ROSECRANS HATHAWAY[1]

Ohio University

INTRODUCTORY

The present paper is divided into three parts. Part One is a statement of the problem and an account of the development and final form of the apparatus. Part Two is devoted to some preliminary findings with the psychogalvanic apparatus comparing reaction times to association word stimuli with reactions obtained by means of this apparatus. Part Three is a general summary with some suggestions as to method of procedure in the use of the psychogalvanic apparatus.

THE APPARATUS

Some four years ago with the idea of studying further the possibilities of the psychogalvanic reflex the apparatus for this purpose described by Wechsler (9) was set up. The findings of our preliminary studies with this apparatus seemed to point to more thorough investigation. Such study was not feasible because of certain difficulties with the method of obtaining and recording the responses. Our main problem, therefore, became that of an attempt to improve the apparatus.

[1] The author wishes to express here his especial gratitude to Dr. James P. Porter particularly whose invaluable interest and criticism made the work possible; to Paul Stocker whose early advice and help laid the foundation for the technical development of the apparatus; to H. P. Longstaff and others on the Ohio University staff who have been personally interested from the first and whose interest and suggestions have meant much.

The particular problems to be met were: first, polarization of the electrodes from the use of direct current made necessary the use of wet or non-polarizing electrodes, thus slowing up the work. The use of a delicate d'Arsonval galvanometer made transportation and setting up a distinctly difficult task.

The first point of attack was the problem of polarization. It was of course recognized that if an alternating current of sufficiently high frequency could be used this difficulty would practically disappear. Work by Galler (10) and Prideaux (7) tended to show that the response could be obtained by the use only of a direct current. Galler used a "pulzenlamp" giving several thousand cycles per second for the extrinsic current and rectified this current after it had passed through the electrodes. Prideaux used a current of 5000 cycles. Both succeeded in eliminating polarization but failed to get responses. It occurred to the writer that these experiments would bear repeating with a somewhat lower frequency.

Accordingly a double cyclometer was built and applied to the Wechsler hook-up in such a way that the first cyclometer changed the direct current from the battery into an alternating current for the bridge and electrodes. The second cyclometer was synchronized with the first and gave this alternating current out again as direct current for the galvanometer which of course could use only direct. The double cyclometer was driven so that about 40 cycles were obtained. To understand this thoroughly the reader should trace carefully through figure 1.

With this arrangement it was found that as expected there was a minimum of polarization and more important the responses were just as strong as with the direct current. This opened the field for further improvement. The cyclometers were noisy and tended to get out of adjustment.

In considering other possible sources of alternating current the ordinary 60 cycle lighting circuit was quite logically the first possibility. This if used would need rectification. The most feasible method for this was the three-element vacuum tube. It was immediately found that if the electrodes were

placed in the grid circuit with the tube arranged in the usual way and that if the galvanometer were placed in the plate circuit an alternating potential of about six volts applied to the grid through the electrodes would show appreciable responses in the galvanometer readings. The discovery of this fundamental circuit led logically to further steps toward greater efficiency. The first move was to put a transformer in the grid

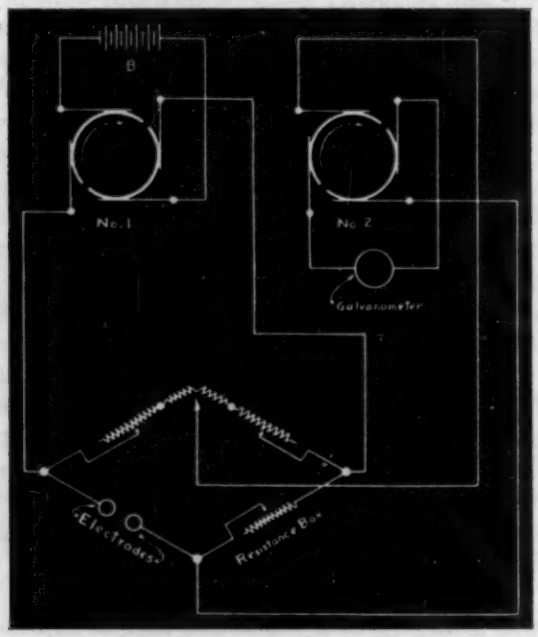


FIG. 1

circuit in such a way that the electrodes were in the primary and the grid with proper C bias was in the secondary circuit. In this way it was possible to obtain a pure alternating current for the electrodes and to allow the use of dry electrodes without fear of polarization.

It was now clear that the responses were capable of much greater amplification. Since current amplification was what

was desired this was accomplished by putting two three-element tubes in parallel. The amplification was so great, thus obtained, that the galvanometer was replaced by an 0–1 milliammeter. One added difficulty was here encountered. There was a constant current of over fifteen times its maximum scale reading flowing in the milliammeter circuit. This was overcome by running an adjustable counter potential across the milliammeter so that the actual current flowing would be nearly zero. The

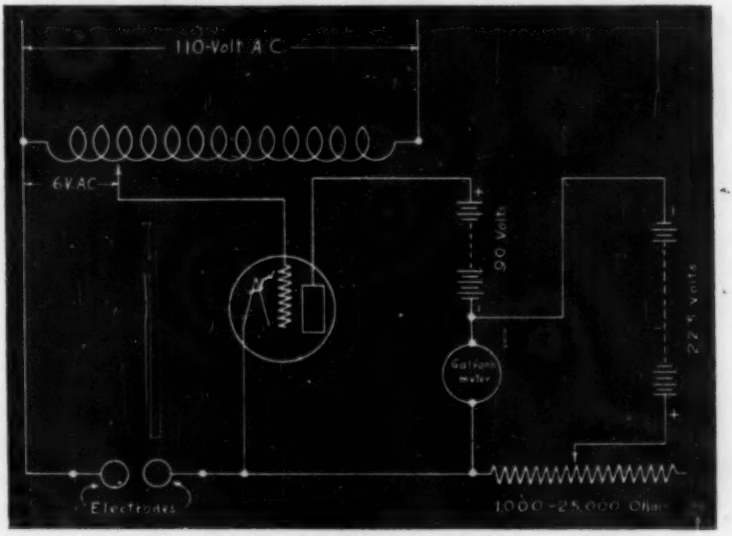


FIG. 2

responses were now read as increments over and above this constant counter potential.

The potential across the electrodes was governed by a potentiometer which gave a voltage range of 0–30 volts. This electrode circuit contained the primary of the transformer through which the grids of the tubes were affected. Negative grid potential was applied by a C battery at 22½ volts. This potential as later trials showed may be kept constant for most uses.

A grid voltmeter was inserted as a means of measuring and keeping constant input potential from the electrodes. The reading on this was dependent upon: (*a*) the value of the C voltage and (*b*) the value of the electrode potential. Since it was desirable to have on every individual subject tested approximately the same electrode current, the second control (*b*) was used, and the voltage was set to a value which depended upon the combined characteristic curves of the tubes used. (The curve for each tube should be nearly the same.)

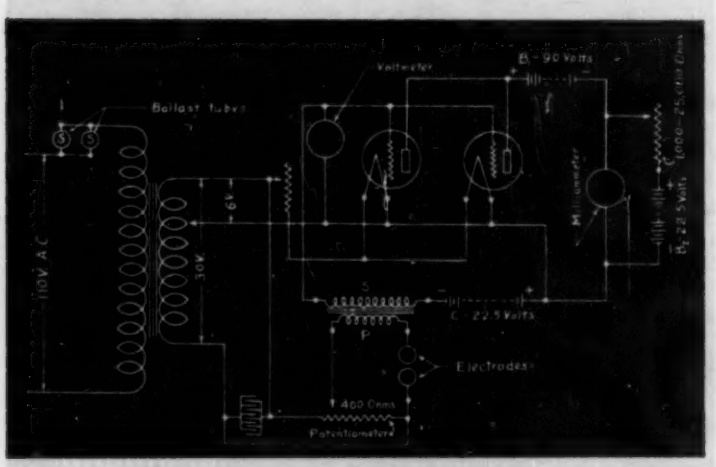


FIG. 3

As the apparatus was finally put together the grid voltmeter was eliminated and advantage was taken of the fact that the grid current varied almost directly with the grid voltage. A double pole double throw switch now served to change the plate milliammeter from the plate circuit to the grid circuit in the preliminary adjustment of the electrode potential. This acted as a substitute for the grid voltmeter. (Fig. 4.)

The electrodes used were of nickeled copper. Two sets were provided. Each electrode of the one set was about 3.5 cm. in diameter and was used in case the subject's hand was dry,

excreting little perspiration. The other set was about 1.7 cm. in diameter and was used on most subjects, or on subjects whose hands were particularly moist. When the subject was in the circuit having the electrodes bound against the palm of his hand the electrode current was adjusted by use of the potentiometer until the reading on the milliammeter in the grid circuit was .01 milliampere or somewhat less depending upon the sensitivity required of the apparatus.

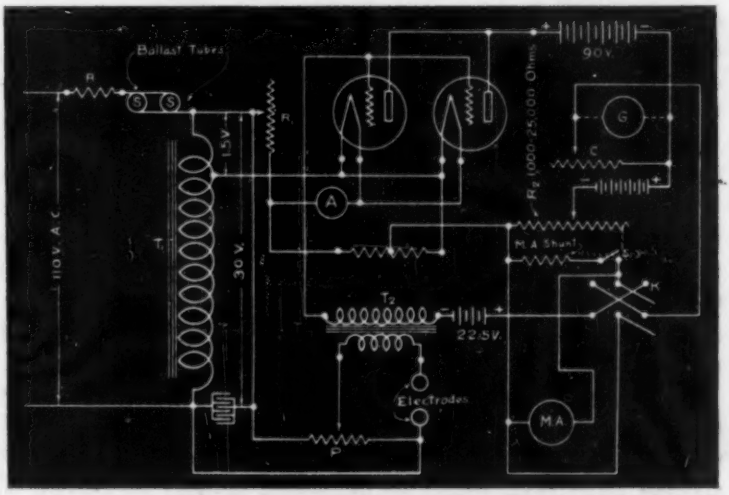


FIG. 4

As will be seen from figure 4 the A.C. 226 type tubes were lighted by alternating current and this supply as well as the alternating current electrode potential was taken from a single transformer giving 1.5 volts and 30 volts respectively. It was found desirable to put two ballast tubes in the primary supply circuit of this transformer so that the variations of line current would not intrude to too great a degree.

In the finished apparatus the batteries, ballast tubes, and transformer were mounted in one case and the controls and meters in another. Though the former proved somewhat

heavy one person could readily carry the complete apparatus. It was found that less than one minute was required to set up the apparatus ready for adjustment of the electrodes to the subject.

The possibility that the resistance of the subject might be an important factor was considered in the construction of the apparatus. In view of the negative results from research on this problem Darrow (1), Syz (2), and in view of the positive fact that the magnitude of the response is proportional to the actual current flowing between the electrodes it was decided that the most efficient method would be to keep this latter current as nearly as possible at a constant value. Accordingly the apparatus was provided with an adjustment for this current. Approximately the same current was used on all subjects tested.

The box containing the controls was supplied also with binding posts, and a shunt control for an external reflecting galvanometer which would be useful for making photographic records and in class demonstrations where it is desired to throw a spot of light on a screen so that the class may observe visual records of the reactions.

The apparatus and method above described measures the reflex responses very quickly and easily. Large numbers of subjects may be tested in a relatively short time. The responses have been found to be strong enough to permit of carrying them over long communicating wires. The apparatus may be placed in one room or building and the electrodes and subject in another. In other words the subject may be isolated and yet his responses accurately measured. The experimenter is thus more completely in control of the stimuli given and of the adequate measurement of the reactions.

PRELIMINARY FINDINGS WITH THE PSYCHOGALVANIC APPARATUS
RESPONSES MEASURED BY REACTION TIME AS COMPARED WITH P.G.R. MEASUREMENTS

The findings herein reported formed a part of a larger problem which was undertaken at the request of the administrative

officers of the university. With this larger problem we are not now chiefly concerned. It is possible, however, to show how the two measures mentioned above compare and to show the effect of diurnal variations on each of these measures. One hundred college women served as subjects. Since all of the subjects knew the purpose of the experiment the averages given below probably are not strictly typical.

An association list of twenty-five stimulus words was selected with two views in mind. First, there were of course certain words the use of which was largely determined by the main problem. Second, words were selected from the Kent-Rosanoff list (5) which were thought to be of particular interest from the

TABLE 1

The twenty-five stimulus words used in the order in which they were given

1. Steal	11. Bills	21. Honor
2. Dark	12. Boy	22. Beautiful
3. Lie	13. Clever	23. Lock
4. Wish	14. Home	24. Tree
5. Dresser	15. Thief	25. Silver
6. Girl	16. Table	
7. Hide	17. Purse	
8. Trouble	18. Justice	
9. Expel	19. Innocent	
10. Dream	20. Afraid	

standpoint of indicating emotional responses. In addition to these several words thought to have little affective value were inserted as a means of control. (Table 1.)

After the electrodes were adjusted the subject was given the usual instructions for the association list. Record was kept of the reaction time, the emotional reaction (the milliammeter reading), and the word which was given as a response. No attempt has been made here to qualitatively analyze the word responses. From subject to subject the word which came first was advanced systematically so that in the one hundred subjects each word in the list came first four times. Otherwise the order was unchanged. A few seconds were allowed between

each word and the next to allow the pointer of the milliammeter to return to its approximate former position.

The averages for the words in tenths of a second of the reaction times and in divisions on the scale on the milliammeter

TABLE 2

The average responses for the twenty-five words in tenths of a second and in hundredths of a milliampere

STIM.	R.T.	P.G.R.
1. Steal	25.6	24.2
2. Dark	22.0	21.0
3. Lie	35.5	25.4
4. Wish	39.6	23.8
5. Dresser	22.9	22.1
6. Girl	23.6	22.9
7. Hide	36.1	23.6
8. Trouble	39.0	27.3
9. Expel	35.2	26.2
10. Dream	30.0	24.7
11. Bills	29.4	23.9
12. Boy	23.3	20.7
13. Clever	31.3	22.5
14. Home	27.4	24.3
15. Thief	29.2	27.8
16. Table	20.8	19.5
17. Purse	23.7	21.7
18. Justice	26.9	26.8
19. Innocent	31.6	24.2
20. Afraid	30.7	24.7
21. Honor	35.6	24.6
22. Beautiful	22.6	21.0
23. Lock	22.5	19.2
24. Tree	23.0	18.0
25. Silver	22.4	22.0

are given in table 2. Each P.G.R. unit represents one one-hundredth of a milliampere. One hundred subjects gave the measures which make up these averages.

The words ranked according to length of reaction time and according to size of milliammeter response are recorded in

separate columns in table 3. Longer reaction times and greater excursion in the milliammeter are assumed to be measures of

TABLE 3

THE TWENTY-FIVE WORDS IN ORDER ACCORDING TO REACTION TIME WITH THE LONG TIMES AT THE TOP OF THE LIST	THE SAME WORDS RANKED ACCORDING TO THEIR RESPONSES ON THE PSYCHOGALVANIC APPARATUS WITH THOSE GIVING THE LARGEST RESPONSES AT THE TOP
1. Wish	1. *Thief
2. *Trouble	2. *Trouble
3. *Hide	3. *Justice
4. *Honor	4. *Expel
5. *Lie	5. *Lie
6. *Expel	6. *Afraid
7. *Innocent	7. Dream
8. *Clever	8. *Honor
9. *Afraid	9. Home
10. Dream	10. *Steal
11. *Bills	11. *Innocent
12. *Thief	12. *Bills
13. Home	13. Wish
14. *Justice	14. *Hide
15. *Steal	15. Girl
16. *Purse	16. *Clever
17. Girl	17. *Dresser
18. Boy	18. *Silver
19. Tree	19. *Purse
20. *Dresser	20. Beautiful
21. Beautiful	21. Dark
22. *Silver	22. Boy
23. Lock	23. Table
24. Dark	24. Lock
25. Table	25. Tree

* Words picked beforehand as having here an especial emotional significance are marked with an asterisk. In the RT list these have an average rank of 18 per cent less than a chance ranking. In the P G R list these have an average rank of 22 per cent less than chance ranking, an increase of 4 per cent in this measure of diagnostic value over the R T ranking. The rank method correlation between these is .60± .086.

stronger emotional disturbance. Those words in the two columns marked with an asterisk were selected beforehand as

likely to arouse the more intense emotions under the conditions involved in this study. It will be noted that the P.G.R. responses seem to be somewhat more diagnostic than the re-

TABLE 4

Total averages for the combined 2500 words in tenths of a second and hundredths of a milliampere

REACTION TIME	PSYCHOGALVANIC RESPONSE
28.4	23.3

Average responses for 625 words given in the morning and 1875 words given in the afternoon

	REACTION TIME	PSYCHOGALVANIC RESPONSE
Morning	27.2	26.0
Afternoon	28.8	22.4

Per cents in morning and afternoon equalling or exceeding the total averages given above

	REACTION TIME	PSYCHOGALVANIC RESPONSE
	per cent	*per cent*
Morning	25.8	48.8
Afternoon	28.4	38.9
Difference	2.6	9.9

Standard errors of difference of these differences computed from Edgerton and Paterson's tables (2)

	REACTION TIME	P.G.R.
Sigma diff.	2.1	2.5
Diff./sigma diff.	1.2	4.5

action times. (See note with table 3.) The most extreme illustration of this is in the case of the word *Thief*. On the association time ranking this comes twelfth while on the P.G.R. ranking it comes first. We should expect this latter to be so.

Some of the other marked words are correspondingly moved upward in the second column.

Particularly interesting is the location of the word *Wish* in the two lists. There seems to be some indication that the long reaction time which is characteristic of this word is due to some other cause than emotion, at least, insofar as the latter is measured by the P.G.R.

In the P.G.R. literature a stimulus word causing marked emotional disturbance has been reported to affect the post critical word. Our results thus far analyzed do not support this conclusion. Illustrating this the word *Table* in the stimulus list followed the word *Thief* yet there is little indication from the rankings that the location of the former word is materially affected by this. At the bottom of this data sheet is given the correlation by the rank method between the two measures. The obtained rho is .60± .086. Using Kelly's table for error of estimate this has a predictability of 20 per cent better than chance.

Table 4 gives first the total averages from the combined two thousand five hundred words for the two measures in the above units. Of the one hundred subjects tested twenty-five were tested in the morning between eight and twelve and the remaining seventy-five in the afternoon between one-thirty and four-thirty. The averages for the six hundred twenty-five words given in the morning as compared with those for the eighteen hundred seventy-five given in the afternoon are also given. In order to obtain some measure of the significance of the differences in these our data were treated in the following manner. A count was taken of the total number of words given in the morning equalling or exceeding the total average and the per cent that this was of six hundred twenty-five was calculated. The same was done with the words given in the afternoon on the basis of eighteen hundred seventy-five. These per cents are given together with their differences in table 4. The standard errors of difference were now computed from Edgerton and Paterson's tables. (2) These latter were now divided into the differences already obtained and show that

in each case the morning and afternoon variations are fairly reliable. The most reliable difference is found in the P.G.R. measures. This difference is four and three-tenths times the standard error of difference. These figures then give a fairly definite indication that: first, the reaction time is typically shorter in the morning, and second, that the P.G.R. reading is typically larger at this same time of day.

We may sum up as follows:

1. The Psychogalvanic apparatus seems to be a somewhat more diagnostic measure than reaction time when association words are used as stimuli to elicit emotional responses.

2. As indicated by our findings the correlation between reaction time measures and P.G.R. readings may be expected to be in the neighborhood of .60 $\pm$.086.

3. Reaction time seems to be typically shorter in the morning.

4. The P.G.R. readings are strongly indicated to be larger in the morning than in the afternoon.

SUMMARY AND SUGGESTIONS AS TO METHODS

The foregoing experiments show with some certainty that our improved apparatus is capable of measuring more readily and clearly some of the aspects of affective behavior. The results are admittedly preliminary and are cited chiefly to show the applicability of this apparatus and the types of problems and methods of procedure to which it lends itself.

Concerning the results on the problem here studied certain points may be quite definitely made. It is rather plainly indicated that this apparatus is somewhat more diagnostic than association reaction time. The diurnal variations in opposite direction of the reaction time and the P.G.R. measures are quite marked and definite.

Farmer (3) and Chambers (4) have called attention to the fact that numerous errors enter into the measurement of the psychogalvanic response *in addition to* those in the apparatus thus making individual subject-to-subject comparison of almost no importance.

There are three possible legitimate methods that are practicable at present. The first is the method of inter-comparison between the stimuli on each subject. This comparison might be in ratio form as in a comparison of the response to success with the same subjects' response to failure (1).

The second method is that of group investigation or comparison of responses such as is used in the experiment with the association word list. The theory behind this is very simple. The additive and subtractive errors will in large part equalize each other.

A third method has been used recently by Wechsler which is even more adapted to the present apparatus. In this records are made of the frequency of response to each stimulus. It seems to the writer that this holds most promise for quantitative comparison of stimuli. Large groups of subjects are used and the frequency of response is recorded without regard to the quantitative value of the single responses. Thus of two stimuli presented to the subjects that one bringing the most frequent response is taken as being the more significant.

In any of these three methods of procedure much of the laborious "standardization of apparatus" which is often indulged in is not only unnecessary but is a waste of time. The only requirements would be of course that the same apparatus be used throughout one series of measurements and that that apparatus be held reasonably near the same sensitivity.

One additional fact is of considerable importance. Even a casual study of the P.G.R. shows it to be very sensitive to the total stimulus situation. It would seem that the matter of first emphasis in studies of this type should be the isolation of stimuli. It is suggested that the general methods of Pavlov in controlling stimulus patterns (6) should be taken into consideration more seriously.

BIBLIOGRAPHY

(1) DARROW, C. W.: Sensory, Secretory, and Electrical Changes in the Skin Following Bodily Excitation. Jour. Exp. Psychol., 1927, Vol. 10, pp. 197–226.

(2) EDGERTON, H. A., AND PATERSON, D. G.: Table of Standard Errors and Probable Errors of Percentages for Varying Numbers of Cases. Jour. Appl. Psychol., Vol. X, No. 3, September, 1926, pp. 378–392.

(3) FARMER, E., AND CHAMBERS, E. G.: Psychogalvanic Reflex in Psychological Experiment. Brit. Jour. Psychol., 1925, Vol. 15, pp. 237.

(4) GEORGI, F.: Beitraege sur Kenntnis des Psychogalvanischen Phanomens. Review by Pieron, H.: L'Année Psychologique, 1922, 23 année, pp. 371.

(5) KENT, C., AND ROSANOFF, A. J.: A Study of Association in Insanity. Amer. Jour. Insanity, 1910, Vol. 67, pp. 48–96.

(6) PAVLOV, I. P.: Conditioned Reflexes, Oxford University Press, 1927.

(7) PRIDEAUX, E.: Expression of Emotions in Cases of Mental Disorder as Shown by the Psychogalvanic Reflex. Brit. Jour. Psychol., 1921.

(8) SYZ, H. C.: Psychological Studies on Sixty-four Medical Students. Brit. Jour. Psychol., 1926, Vol. 17, No. 27, pp. 54–69.

(9) WECHSLER, I. S.: The Measurement of Emotional Reactions. Arch. Psychol., No. 76.

(10) Quoted from (9), p. 102.

(11) YERKES, R. M. The Study of Human Behavior, Science, 1914, Vol. 39, pp. 625–633.

NEWS AND NOTES

Exposition de Liege, 1930: During the year 1930 at Liege the people of Belgium will celebrate by means of an international exposition the century of their national independence. The plans provide a special place for the exhibit of diagrams, synoptic charts and tables, apparatus, teaching materials, tests, photographs and questionnaires relative to scientific organization, selection, placement and other personnel problems of modern business, industry and education. Those interested in this most promising and interesting section should address Mr. Lucien Wellens, Secretaire de la Classe, 94 D, 11, rue de Seraing, Liege, Belgium.

The Tenth Annual Ohio State Educational Conference: "Reaching the Individual" will be the keynote of the Tenth Annual Ohio State Educational Conference to be held in Columbus, April 3, 4, 5, 1930. Mr. Robert M. Hutchins, president of the University of Chicago, will speak at the Thursday night general session. Mr. E. H. Sothern, well-known actor and dramatic reader, will give a series of readings from Shakespeare Friday night. One additional speaker for the Thursday night general session is yet to be secured.

More than one hundred speakers, including some thirty-five from out of the state, will participate in the Tenth Annual Ohio State Educational Conference in which a registration of over five thousand is expected. Last year 5100 were registered.

Visual education will be a new topic for which sectional meetings will be held. This brings the total number of sections to thirty-five. Three sessions of conference (Friday morning and afternoon and Saturday morning) will be given over to these sectional conferences.

Dr. J. E. W. Wallin, who for eight years was Director of the Bureau of Special Education and the Psycho-Educational Clinic and Professor of Clinical Psychology in Miami University, is now Director of the Division of Special Education in the Baltimore Department of Education which was established following a survey of special classes in Baltimore which he made last spring. Dr. Wallin also offers courses on the education and clinical study of handicapped children at Johns Hopkins University and Morgan College. He will offer similar courses in the first summer term at Johns Hopkins and in the second summer term at the University of Chicago.

BOOK REVIEWS

J. L. Prak. *De Psychotechniek der Beroepskeuze.* J. B. Wolters' U. M., Groningen, Den Haag, 1925. £4.50.

This is a critical compilation work by a Dutch psychologist, who is in charge of the applied psychological work in the Philips Factories. There is very little that will prove to be new to the psychologist who is well acquainted with the literature on Employment Psychology and Vocational Selection. This book is however an excellent introduction into this field for the layman, or for the psychologist who is not in touch with this branch of his subject. It therefore serves the purpose the author set himself, viz.: to give a survey of the field and problems of the psychology of vocational and personnel selection; to demonstrate, with proofs gathered from the investigations of many authors, the value of experimental research methods and precise statistical methods of calculation in connection with such selection; and to warn against the too optimistic belief in so-called intuitive knowledge of mankind. The author draws very heavily from many sources and the Bibliography succeeding each chapter contains references to Dutch, French, English, American and German works. The latter may prove useful to those being introduced to the subject but contains nothing unknown to the worker in the field. A brief summary of the contents will show to what extent the book is a compilation.

In Chapter 1 the factors which at present influence the choice of an occupation are described in the usual way and the ordinary methods of selecting a job are criticised. Opinions of well-known people and American data on labour turnover are given. Chapter 2 contains a discussion of the reasons for their choice of an occupation given by schoolboys and students according to the investigations of Minkus, Jonckheere, Fischer and especially of Wisse whose work carried out in Holland is cited at some length to show that the student does not choose a profession but a subject for study, and that ultimately the issue is decided not by the demands of the profession or the qualities of the individual but by financial attractiveness of the position, family tradition, etc. Hollingworth is cited at length in the discussion of the value of self-estimates, and also, together with Rugg and the Army Rating Scale, in connection with the value of the rating of characteristics by others. Roloff's results on the value of teachers' judgments of their schoolchildren's intelligence are also given. The value of school attainments

as an indication of vocational aptitude is discussed in connection with the results of Kelley, Miles, Dearborn, Smith and Thorndike. Hollingworth is also the chief source from which the material for Chapter 3 is drawn which contains a criticism of the usual methods for choosing personnel. In connection with the value of letters of application Poffenberger and Vartanian are also cited, in connection with the value of photographs also Pintner, Anderson, Cleaton and Knight, and in connection with the personal interview also Binet, Scott and Snow. Chapter 4 contains an extremely brief sketch of the development and subject-matter of Psychology, and Chapter 5 a short survey of the procedure in Vocational and Employment Psychology. In Chapter 6 the following are described at some length as a few of the chief events in its history: Münsterberg's work and experiments, the Binet-Simon tests, the American Army tests, Moede and Piorkowski's tests for the Begabtenschule in Berlin, and Stern's Hamburg tests for a similar purpose. Chapter 7 chiefly contains illustrations of investigations of several kinds of occupations, e.g., Tramm's of Tram-drivers; that of the German Railway's investigation Bureau of apprentices in the workshops; that of Giese, as well as that of Brugmans and Prak, of Exchange Girls; Cyril Burt's of Stenographers; and Schulte's of the Ladies' Hairdresser. In Chapter 8 the present position of Vocational Psychology in the different countries is described, especially in the Netherlands. Objections against this branch of science are stated and discussed in Chapter 9; and in Chapter 10 the author criticises very sharply the standpoint of those who favour the use of various types of questionnaires instead of experimental investigations in connection with the problems of selection. Chapter 11 is an elementary one on the statistical methods used, and in Chapter 12 the education of the practical psychologist is briefly discussed in connection with a diploma course of London University, and the different types of institutions in which Psychotechnic is developing are mentioned.

The book is, however, not just a compilation but critical throughout and reveals the author's scientific attitude and practical experience. There is no over-enthusiasm but the exposition of facts in a sound and sensible manner. This makes it all the more valuable as an introductory work, especially as the attempt has evidently been made to be as clear as possible by just presenting the generally accepted outline of what is and can be done, and not entering to any extent upon difficulties which still perplex the specialist. The fact must nevertheless be recognized that even in the critical remarks on the whole the initiated will find nothing very striking or new, so that one does not feel that anything has been added to one's usual point of view. An exception is formed by Chapter 10 and anyone interested in making use of questionnaires in connection with Vocational Selection or Employment Psychology, or

wishing to obtain answers from teachers to lists of questions on the characteristics of their scholars for some other purpose, would do well to read this chapter.

H. F. VERWOERD,
University of Stellenbosch.

GAST AND SKINNER. *Fundamentals of Educational Psychology.* Benj. H. Sanborn Co., New York City, 1929. 354 pp.

The authors have departed from the usual text in educational psychology in several ways. One of the improvements is the inclusion of a chapter on childhood and adolescence. In this chapter they discuss problems of growth and development which are usually excluded from educational psychologies. Criticisms of this chapter are undoubtedly due because of its brevity and lack of data, but the authors do deserve commendation for including it at all.

Two other departures are somewhat interesting. One is the inclusion of about two score pictures of psychologists and educators—many of them living. Much is made of psychological authorities, and the book bristles with "Thorndike says, Burt states, Burnham reminds us," etc. It seems to the reviewer that the student is apt to become a worshipper of authority instead of a critical thinker dealing with data. Another departure from the usual text was the failure to make any attempt to provide a logical system of educational psychology. What advantage there can be in an unsystematic treatment the reviewer fails to see. Certainly the laws of learning and retention make it clear that well organized material will be retained more easily and longer—but perhaps a psychologist is not expected to observe the laws of learning.

STUART M. STOKE,
Ohio University.

EDMUND S. CONKLIN. *The Psychology of Religious Adjustment.* The Macmillan Company, New York, 1929. Pp. xiv + 340.

As the title suggests the author thinks of religion as a method of adjustment. He makes a good case for this particular way of regarding it without letting his thesis so influence his material and treatment as to greatly impair the value of his work for those who regard religion in some other way: as divine revelation, infantile regression, or what not. This concept does, however, color the author's explanations of certain religious phenomena and it is in these explanations that most of the originality the work may possess is to be found.

It may not be amiss to point out a few features of the book which seem especially commendable. Many students of the psychology of religion fall short of the scientific ideal in that they are either friends or enemies of religion and clearly show the effect of their attitudes in their

writing. In the opinion of the reviewer the author has avoided this error to a greater degree than have the majority of writers in this field. It is also to be noted that the psychology of conversion is given a less prominent place than it has traditionally held, being discussed adequately in one chapter of medium length. This is especially commendable in view of the decreasing emphasis on conversion in many religious sects. Ecstatic states including "power states," prophecy and mysticism are treated in two chapters, and religious exercises in three. Much of the value of the book consists in the treatment of certain phenomena often neglected. Sanctification, the retreat, speaking with tongues, and apocalypticism may be mentioned as examples.

The author has little sympathy for the emphasis often placed on the sex element. That the love of God is derived from the primitive sex instinct, that mysticism may be traceable to sexual abnormalities he refutes. For McDougall's love sentiment he finds a place; for the interpretation of the Freudians, none.

AMOS C. ANDERSON
Ohio University

NEW BOOKS AND PAMPHLETS*

Books and pamphlets for review should be sent to James P. Porter, Department of Psychology, Ohio University, Athens, Ohio.

American Hand Writing Scale. PAUL V. WEST. The A. N. Palmer Company, New York City.

Characteristic Differences of Good and Poor Teachers. A. S. BARR. Public School Publishing Company, Bloomington, Ill. 127 pp.

Conquest of Thought by Invention, The. H. STAFFORD HATFIELD. W. W. Norton & Company. $1.00. 80 pp.

Crime and the Press. JOSEPH L. HOLMES. From Journal of Criminal Law and Criminology, Vol. XX, Nos. 1 & 2. 1929. 107 pp.

Diagnostic and Remedial Work in Arithmetic Fundamentals. MAY LAZAR. Board of Education, New York City. Bureau of References, Research and Statistics. 204 pp.

Ethics. FRANK C. SHARP. The Century Company, New York City. $3.50.

Foundations of Experimental Psychology, The. CARL MURCHISON, Editor, and twenty-three contributing authors. Clark University Press, Worcester, Mass. 1929. $4.00. 907 pp.

Fundamentals of Educational Psychology. IRA MORRIS GAST AND HARLEY CLAY SKINNER. Benjamin H. Sanborn & Company, New York City. 354 pp.

Future of the Earth, The. HAROLD JEFFREYS. W. W. Norton & Company. $1.00. 79 pp.

How to Find the Right Vocation. HARRY DEXTER KITSON. Harper & Brothers, New York City. $2.50. 202 pp.

Introduction to Abnormal Psychology. V. E. FISHER. Macmillan Company, New York City. 512 pp.

Just Normal Children. FLORENCE MATEER. D. Appleton & Company. $2.00. 294 pp.

Musician's Arithmetic, The. MAX F. MEYER. University of Missouri Studies, Vol. IV, No. 1, Columbia, Mo. 149 pp.

Measurement of Attitude, The. L. L. THURSTONE AND E. J. CHAVE. University of Chicago Press, Chicago, Ill. $1.00. 96 pp.

Mind at Mischief, The. WILLIAM S. SADLER. Funk & Wagnalls Company, New York City. $4.00. 400 pp.

---

* Mention here does not preclude further comment.

Newer Ways with Children. M. V. O'Shea. Greenberg, Publisher, New York City. 419 pp.

An Outline of Social Psychology. J. R. Kantor. Follett Publishing Company, Chicago, Ill. 420 pp.

Practical Salesmanship. W. L. Barnhart. The Ronald Press Company, New York City. $3.50. 266 pp.

Psychological Register, The. Carl Murchison and others. Clark University Press, Worcester, Mass. $6.00. 580 pp.

Psychology of Adolescence, The. Fowler D. Brooks. Houghton Mifflin Company. 652 pp.

Psychology. Revised Edition. Robert S. Woodworth. Henry Holt & Company, New York City. $3.00. 590 pp.

Relative Frequency as a Determinant of Phonetic Change. George Kingsley Zipf. From Harvard Studies in Classical Philology, Vol. XL, 1929. 95 pp.

Science of Psychology. Raymond Holder Wheeler. Thomas Y. Crowell, New York City. $3.75. 556 pp.

Social Psychology. Bernard C. Ewer. Macmillan Company, New York City. 436 pp.

Social Psychology of International Conduct. George Malcolm Stratton. D. Appleton Company, New York City. $3.00. 387 pp.

Sociology and Sin. P. Sargant Florence. W. W. Norton & Company, New York City. $1.00. 98 pp.

Study Manual in Elementary Statistics. Everett F. Lindquist and George D. Stoddard. Longmans, Green & Company, New York City. $1.80. 109 pp.

Survey of Special Education in the Public Schools of Baltimore. J. E. W. Wallin. Department of Education, Baltimore, Md. 57 pp.

Testing Intelligence and Achievement. Albert J. Levine and Louis Marks. Macmillan Company, New York City. 399 pp.

INDEX OF SUBJECTS

Titles of books are followed by the name of the author of the work reviewed. All other references are original articles.

Abilities, young people, differences, talents and, 451.

Ability and achievement in general psychology, objective measures, 173; —— groups, measurement, reading skills, 494; ——, prognosis, sight singing, 425.

Accomplishment quotient, correlation, intelligence quotient and, 76.

Accountants and bookkeepers, predictive value, intelligence tests, selection of junior, 1.

Achievement in general psychology, objective measures of ability and, 173.

Action on voice intensity, effect of bodily, 516.

Activities, investigation of some simple speed, 167.

Adult-level norms, personnel procedure, interpolated revision, college, 325.

Advertising appeals, measured by coupons returned, relative values of positive and negative, 274; ——, historical trend, negative appeals, 346; ——, memory value, absolute size, magazine, 62.

Age, improvement, feeble-minded girls over sixteen years, 301.

Age-level, mental alertness examination, working, 46.

Agriculture, measurement of interest differences, students of engineering and, 105.

Anderson tests, Pearson r and further note on Kuhlmann, 32.

Apparatus, new psychogalvanic, 632

Appeals in advertising, historical trend, negative, 346; ——, measured by coupons returned, relative values, positive and negative advertising, 274.

Aptitude testing (Hull), 95.

Aptitude tests, what measured by Iowa, 589

Armenian, Turkish, Greek, Bulgarian subjects, mental test findings, 266.

Association time measures, comparative study of psychogalvanic and, 632

Athletics and scholarship (Paterson and Peterson), 322.

Australian aboriginal, social psychology, 131.

Bookkeepers, predictive value, interest tests, selection of junior accountants and, 1.

Bulgarian, Armenian, Turkish, Greek subjects, mental test findings, 266.

Children, manual motor correlation, superior, 357

Children's behavior and teachers' attitudes (Wickman), 323.
Child's picture and story books, battery of tests method, preschool, 592.
Chinese and Japanese, mentality of, 9.
Clinical and abnormal psychology (J. E. W. Wallin), 94.
College adult-level norms, personnel procedure, interpolated revision, 325; —— freshmen, avocations of one hundred, 256; —— student, maladjusted, 440.
Context, relative effects, size of type, leading and, 600.
Corrected, special disability in spelling, diagnosed and, 378.
Correlation, superior children, manual motor, 357.
Coupons returned, relative values, positive and negative advertising appeals, measured by, 274.
Criminal intelligence (Murchison), 96.
Criterion, predictive value, mental tests that satisfy Spearman's tetrad, 152.

Deaf, speech and speech-reading tests, 220.
Delinquency, Porteus maze test as discriminative measure, 145.
Diagnosed and corrected special disability, spelling, 578
Difference in handwriting, sex, 159
Differences, students of engineering and of agriculture, measurement of interest, 105.
Differential IQ's of Siblings, 86.
Disability, spelling, diagnosed and corrected, special, 378.
Educational Psychology (Gast and Skinner), 650.
Educational Psychology (Pintner), 539.
Educational Psychology (Sandiford), 198.
Effects of music, the (Schoen), 199.
Emotional instability, objective measures, 226.
Employability, standardized emotional tests, validity and, 554.
Engineering and agriculture, measurement of interest differences between students, 105.
Examination, working age level, mental alertness, 46.
Extroversion, scale, measuring introversion-, 499.

Factors influencing speed and reading, studies of typographical, 205.
Feebleminded girls over sixteen years of age, improvement, 301.
Freshmen, avocations of one hundred college, 257.

General psychology, objective measurements, ability and achievement in, 173.
Girls over sixteen years of age, improvement, feebleminded, 301.
Greek, Bulgarian, Armenian, Turkish subjects, mental test findings, 266.

Habit modification, conceptual basis, 469.
Handwriting, sex difference, 159.
History of psychology, the (Pillsbury), 535.
Human nature and management (Tead), 536.

Infancy and Human Growth (Gesell), 321.
Instability, objective measures, emotion, 226.
Intelligence quotient and accomplishment quotient, correlation, 76; —— tests, comparative validity, primary, 394; —— tests, selection of junior accountants and bookkeepers, predictive value of general, 1; ——, validities of judgments, 630.
IQ's of siblings, differential, 86.
Intensity, effect, bodily action on voice, 516.
Interest differences, students of engineering and of agriculture, measurement, 105.
Introduction to modern psychology (Murphy), 421.
Introversion-extroversion, scale, measuring, 499.
Iowa aptitude tests, what measured, 589.

Japanese, mentality of Chinese and, 9.
Judgment, errors, self, 372.
Judgments, intelligence, validities, 630.

Kuhlmann-Anderson tests, Pearson r and further note, 32.
Kuhlmann-Binet tests, the (Good-enough), 94.

Laboratory exercises in educational statistics (Morton), 417.
Leading and context, relative effects, size of type, 600.
Line, length of, studies of typographical factors influencing speed and reading, 205.
Lispers, experiment with, 543.

Magazine advertising, memory value, absolute size, 62.
Marketing and advertising (Vaughan), 195.
Maze test as discriminative measure in delinquency, Porteus, 145.
Measure of emotional instability, objective, 226.
Measures of ability and achievement in general psychology, objective, 173; ——, comparative study, psychogalvanic and association time, 632.
Mental tests that satisfy Spearman's tetrad criterion, predictive value, 152.
Method, preschool child's picture and story books, battery of test, 592.
Modification, conceptual basis, habit, 469.
Motor correlation, superior children, manual, 357.
Musical talent, selecting, 486.

Norms, personnel procedure, interpolated revision of college adult-level, 325.

People, differences in talent and abilities, young, 451.
Personnel procedure, interpolated revision, college adult-level norms, 325.
Picture and story books, battery of tests method, preschool child's, 592.
Preschool child's picture and story books, battery of tests method, 592.
Primary intelligence tests, comparative validity, 394.
Prinzipien und methoden der kunstpsychologie (Plaut), 419.
Prognosis, school success in typewriting, study of, 505.

Psychiatry in industry (Anderson), 418.
Psychogalvanic apparatus, new, 632; —— and association time measures, comparative study, 632.
Psychological and educational survey of 1916 prisoners in the Western Penitentiary of Pennsylvania (Root), 96.
Psychology, experiment in teaching, 188; ——— of the Australian aboriginal, social, 131; ——, objective measures of ability and achievement in general, 173; —— test, college, 81.
Psychology of Religion (Conklin), 650.

Race attitudes in children (Lasker), 320.
Reading skills in ability groups, measurement, 494; —— studies of typographical factors influencing speed, 120; —— studies of typographical factors influencing speed, (length of line), 205; —— tests for the deaf, speech-, 220.

Scale, measuring introversion-extroversion, short, 499.
School success, typewriting, prognosis, 505.
Selection, junior accountants and bookkeepers, predictive value, general intelligence tests, 1.
Self judgment, errors, 372.
Siblings, differential IQ's, 86.
Sight singing ability, prognosis, 425.
Singing ability, prognosis, sight, 425.
Size in magazine advertising, memory value, absolute, 62.
Size of type, leading and context, relative effects, 600.
Skills, ability groups, measurement of reading, 494.
Spearman's tetrad criterion, predictive value, mental tests that satisfy, 152.
Special disability, spelling, diagnosed and corrected, 378.
Speech-reading tests for the deaf, speech and, 220.
Speed activities, investigation of some simple, 167.
Speed of reading, studies of typographical factors influencing, 205; —— of reading, (size of type), 120.
Spelling, diagnosed and corrected, special disability, 378.
Standardized emotional tests, validity and employability, 554.
Story books, battery of tests method, preschool child's picture and, 592.
Student, maladjusted, college, 440.
Students of engineering and agriculture, measurement of interest differences, 105.
Study of educational achievement of problem children (Paynter and Blanchard), 537; —— of psychogalvanic and association time measures, comparative, 632; —— with results, maladjusted college student, 440.
Subjects, mental test findings, Armenian, Turkish, Greek, Bulgarian, 266.
Success in typewriting, prognosis of school, 505.
Superior children, manual motor correlation, 357.

Talent, selecting musical, 486.

Talents and abilities, young people, differences, 451.
Teaching psychology, experiment, 188.
Test as discriminative measure in delinquency, Porteus maze, 145; ——, college psychology, 81; —— findings, Armenian, Turkish, Bulgarian, Greek subjects, mental, 266.
Tests, comparative validity, primary intelligence, 394; —— for the deaf, speech and speech-reading, 220; ——, measured, Iowa aptitude, 589; —— method, preschool child's picture and story books, battery of, 592; ——, Pearson r and further note on the Kuhmann-Anderson, 32; ——, selection, Junior accountants and bookkeepers, predictive value of general intelligence, 1; —— that satisfy Spearman's tetrad criterion, predictive value of mental, 152; ——, validity and employability, standardized emotional, 554.
Tetrad criterion, predictive value of mental tests that satisfy Spearman's, 152.
Turkish, Greek, Bulgarian, Armenian subjects, mental test findings, 266.
Type, leading and context, relative size, 600; ——, size, studies of typographical factors influencing speed of reading, 120.
Typewriting, prognosis, school success, 505.
Typographical factors influencing speed, reading, studies, 120, 205.

Validity and employability, standardized emotional tests, 554; —— primary intelligence tests, comparative, 394.
Voice intensity, effect, bodily action, 516.

Whither mankind: a panarama of modern civilization (Beard), 196.
Working age level, mental alertness examination, 46.

INDEX OF AUTHORS

The names of authors of original contributions are printed in CAPITALS AND SMALL CAPITALS.

Anderson, Amos C., 94, 196, 321, 536, 651.
ASHER, E. J., 152.
BALL, ROBERT JAUDON, 226.
BATHURST, J. E., 81.
BENSON, C. E., 274, 346.
BOOK, WILLIAM F., 378.
BOUTON, M. THELMA, 159.
BROOM, M. EUSTACE, 159, 325.
Burtt, Harold E., 203, 421, 537.
BERRY, CHARLES SCOTT, 543.
CLINE, W. F., 257.
DOUGLAS, JOSEPHINE, 494.
DOUGLASS, HARL R., 76.
Fisher, Sara C., 102.
GARRISON, KARL C., 167.
Gentry, J. R., 538.
GILLESPIE, F. RODNEY, 325.
GORHAM, DONALD R., 554.
HATHAWAY, STARKE ROSECRANS, 632.
HEIM, M. H., 62.
HIGHSMITH, J. A., 486.
HOVDE, HOWARD T., 600.
HSIAO, HSIAO HUNG, 9.
HUFFAKER, C. L., 76.
JACKSON, T. A., 372.
KEIFER, FRIEDA A., 357.
KUHLMANN, F., 32.
LANGLIE, T. A., 589.
LAWSON, J. W., 494.
Lehman, Harvey C., 198.
LONGSTAFF, HOWARD P., 173
LUCAS, D. B., 274, 346.
McFADDEN, J. H., 86.
MILLER, ELEANOR OLMSTEAD, 592.
MONROE, ALAN H., 516.
MONTGOMERY, RUTH P., 145.
MORIWAKI, ETHEL, 630.
NATHANSON, YALE S., 469.
NEWHALL, S. M., 62.
OTIS, MARGARET, 301.
PATERSON, DONALD G., 120, 205.
Patrick, James R., 322, 422.
PINTNER, RUDOLF, 220.
PORTER, JAMES P., 173.
PORTEUS, S. D., 131.
POULL, LOUISE E., 145.
REMMERS, H. H., 105.
SALISBURY, FRANK S., 425.
SANGREN, PAUL V., 394.
SCHEIDEMANN, NORMA V., 81, 188.
SCHULTZ, RICHARD S., 46.
SCUDDER, K. J., 1.
SMITH, H. B., 425.
STEDMAN, MELISSA BRANSON, 505.
STODDARD, CLARA B., 543.
STOGDILL, EMILY LEATHERMAN, 440.
STOKE, STUART M., 95, 96, 199, 257, 323, 418, 539, 650.
THOMPSON, BLANCHE, 159.
TINKER, MILES A., 120, 205.
TYBJERG, CHR. HANSEN, 451.
Verwoerd, H. F., 650.
WHITMAN, ROSWELL H., 499.
WOOD, MARGARET M., 266.

Directory of American Psychological Periodicals (*Continued*)

Journal of Personnel Research—Baltimore: Williams & Wilkins Co. Subscription $5.00. 500 pages approximately. Edited by Walter V. Bingham, and Clarence S. Yoakem. Bi-Monthly.

Genetic Psychology Monographs—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison and an international co-operating board. Monthly. Each number one complete research. Child behavior, differential and genetic psychology. Founded 1925.

Psychological Abstracts—Princeton, N. J. Subscription $6.00. 600 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

Mental Measurement Monographs—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Founded 1925. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research.

Studies in Psychology and Psychiatry—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Founded 1926. Edited by Edward A. Pace. Published without fixed dates, each number a single research.

Journal of General Psychology—Worcester, Mass.: Clark University Press. Subscription $7. 600–700 pages annually. Edited by Carl Murchison and an international co-operating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Directory of American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Morrill Hall.
Subscription $6.50. 600 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.: Clark University Press. Subscription $7. 700 pages annually. Edited by Carl Murchison and an international co-operating board. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 480 pages annually. Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

Psychological Bulletin—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Founded 1904. Edited by Samuel W. Fernberger.

Psychological Monographs—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Founded 1895. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300–400 pp. Founded 1895. Edited by W. S. Hunter. An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116th St.
Subscription $4. 728 pages per volume. Founded 1904.
Bi-weekly. Edited by F. J. E. Woodbridge and Wendell T. Bush.

Archives of Psychology—Sub-Station 84, N. Y.: Archives of Psychology.
Subscription $6. 500 pp. ann. Founded 1906. Ed. by R. S. Woodworth.
Published without fixed dates, each number a single experimental study.

Journal of Abnormal and Social Psychology—Eno Hall, Princeton, N. J.
Subscription $5. Foreign, $5.25. Acting Editor, Henry T. Moore.
Quarterly. 432 pages annually. Founded 1906. Abnormal and Social.

Psychological Clinic—Philadelphia: Psychological Clinic Press.
Subscription $2.50. 288 pages. Ed. by Lightner Witmer. Founded 1907.
Without fixed dates (9 Numbers). Orthogenics, psychology, hygiene.

Training School Bulletin—Vineland, N. J.: The Training School.
Subscription $1. 160 pp. ann. Ed. by E. R. Johnstone. Founded 1904.
Monthly (10 numbers). Psychology and training of defectives.

Comparative Psychology Monographs—Baltimore: Johns Hopkins University Press. Subscription $5. 500 pages per volume. Edited by W. S. Hunter.
Published without fixed dates, each number a single research.

Psychoanalytic Review—Washington, D. C.: 3617 10th St., N. W.
Subscription $6. 500 pages annually. Psychoanalysis.
Quarterly. Founded 1913. Ed. by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental.
Subscription $6.00. Founded 1916. Bi-monthly. Ed. by Madison Bentley.

Journal of Applied Psychology—Baltimore: The Williams & Wilkins Co.
Subscription $5.00. 400 pages annually. Founded 1917. Bi-monthly.
Edited by James P. Porter, Ohio University, Athens, Ohio.

Journal of Comparative Psychology—Baltimore: Williams & Wilkins Co.
Subscription $5. 500 pages annually. Founded 1921.
Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

(*Continued on Third Cover*)